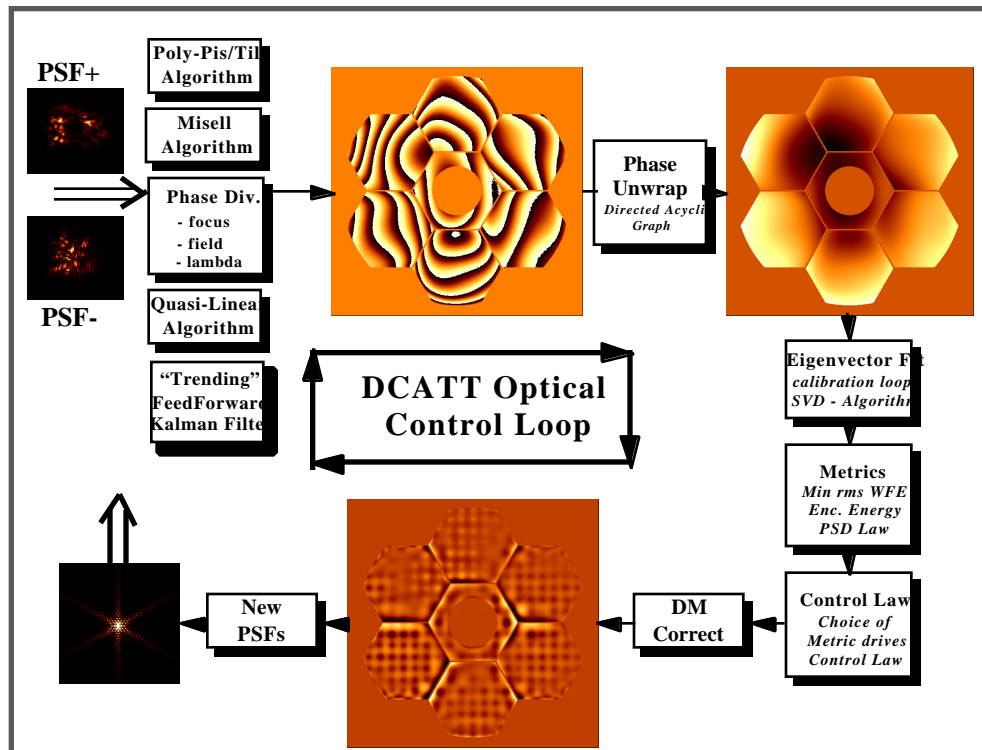# DCATT Wavefront Sensing and Optical Control Study

**Developmental Comparative Active Telescope Testbed
Report WFSC-0001
February 22,1999**

This report is submitted January 15th, 1999 in partial fulfillment of NASA Contract NAS5-32337 Task No. 57 Mod E, initiated December 3rd, 1997.

Author: Richard G. Lyon   NASA/GSFC/CESDIS Code 930.5
                                        301-286-4302
                                        lyon@jansky.gsfc.nasa.gov

Contributors: Timothy P. Murphy   NASA/GSFC/CESDIS Code 930.5
                                        301-286-9805
                                        murphy@albert.gsfc.nasa.gov

                 John E. Dorband   NASA/GSFC Code 935.0
                                        301-286-9419
                                        dorband@newton.gsfc.nasa.gov

                 Eric W. Young   NASA/GSFC Code 551.0
                                        301-286-1366
                                        Eric.W.Young.1@gsfc.nasa.gov

                 Mark E. Wilson   NASA/GSFC Code 551.0
                                        301-286-5538
                                        Mark.E.Wilson.1@gsfc.nasa.gov

                 Pamela S. Davila   NASA/GSFC Code 551.0
                                        301-286-2074
                                        Pamela.S.Davila.1@gsfc.nasa.gov

Abstract

The Developmental Comparative Active Telescope Testbed (DCATT) is a ~1 meter, segmented aperture, optical system consisting of a telescope, active optical bench, source and focal plane. DCATT is currently under development at the NASA/Goddard Space Flight Center (GSFC) in Greenbelt, MD, as a collaborative effort between GSFC and the Jet Propulsion Laboratory (JPL). The primary purpose is to represent, in hardware, an optical system which incorporates the functionality of an NGST active optical system. Also, to investigate the various methods of wavefront sensing and optical control, initially to bring the system into alignment (coarse alignment) and subsequently maintain alignment (fine figure control) at diffraction limited performance at 2 microns. The ultimate goal is to utilize the information, obtained from the DCATT experiments, in the design of the wavefront sensing and optical control system for the Next Generation Space Telescope (NGST). A further objective is to develop and validate high fidelity integrated computer models of the DCATT system, thus, gaining confidence in our ability to model complex optical

systems, and, predict their performance. This is an important consideration in the design of NGST and other future NASA flight missions.

We primarily report on the results of our study with regards to fine figure wavefront sensing and optical control. An detailed description of the coarse alignment process is reported by Redding et.al. [1] and lessons learned from Hubble Space Telescope is reported on in [2]. This study is comparative in both a qualitative and quantitative sense and compares phase retrieval and Shack-Hartmann sensing, and discusses number of different optical control loop options using different metrics. This study is not meant to be definitive since more work is required. This will be discussed in further detail in section 6.

Table of Contents

1.0 Executive Summary

This report details the results on an on-going comparative wavefront sensing study for DCATT. In this study we compare, through the use of computer and analytical models, a focus diverse (Misell algorithm) phase retrieval algorithm, a wavelength diverse phase retrieval algorithm, an extended object phase diversity algorithm, and a Shack-Hartmann sensor. We rank the limited subset of errors addressed here in terms of importance and make recommendations on what methods and techniques need further study, to possibly improve wavefront sensing for DCATT. We also discuss, model and simulate some optical control loop options. This study in entirely based on high fidelity computer modeling and simulations of DCATT. Our models are meant to be predictive in nature of DCATT's actual performance and will ultimately be adjusted and updated as needed when DCATT comes online at NASA/GSFC in March 1999. It is expected that at some point in late 1999 or early 2000 that the models and DCATT observed behaviors will agree, thereby, gaining confidence in this and other models for future use on other missions such as NGST.

The overall layout of this report is in 5 parts: (i.) an introduction, discussing the relevance of DCATT in the design of NGST, (ii.) a discussion of the optical modeling techniques used, (iii.) a comprehensive description of both the phase retrieval and Shack-Hartmann wavefront sensing studies performed, (iv.), a discussion of optical control and calibration techniques and (v.) recommendations based upon this study. Note a detailed list of references is included for each section of this report.

2.0 Introduction

Currently four study teams (GSFC/JPL, Lockheed-Martin, TRW and Ball) have developed conceptual designs and mission architectures for the Next Generation Space Telescope (NGST). See references [3][4][5] for a more detailed description of NGST and it's science requirements. The 4 groups presented a number of different concepts for wavefront sensing (WFS) and optical control. In general, the techniques fell into three broad categories of wavefront

sensing: (1) phase retrieval techniques, (2) Shack-Hartmann techniques, and (3) interferometric techniques. Phase retrieval is the most computationally complex method of the three but does not require any additional hardware and uses one or more science cameras. Both Shack-Hartmann and interferometry require more hardware and calibration, however, but are significantly less computationally complex than phase retrieval. All these techniques have been previously demonstrated on monolithic aperture optical systems. One challenge for NGST is to demonstrate that these methods will also work, in a closed loop system, operating at reduced data rates and at cryogenic temperatures in a space environment. The method chosen for NGST must ultimately be capable of measuring and controlling the wavefront to the required accuracy, precision and repeatability at low cost and high reliability.

NASA is currently developing a series of ground testbeds and validation flights to demonstrate the technologies of deployable, lightweight optics, cryogenic optics and mechanisms, passive, cooling, and on-orbit, closed loop wavefront control techniques. The Developmental Comparative Active Telescope Testbed (DCATT) [6][7][8][9] is a 1 meter class system level testbed currently under development at Goddard Space Flight Center in collaboration with the Jet Propulsion Lab. This testbed combines a segmented aperture telescope with the hardware elements of an active optical bench and the software elements of an active control system to demonstrate diffraction limited performance at 2 microns. Although a baseline wavefront sensing and control scheme for DCATT has been chosen, we need to better understand the advantages and disadvantages of all the 3 aforementioned techniques so that we may be better qualified to select, and optimize, the best technique for NGST. This study represents one step in that process. It is our hope that one or more techniques for each of the 3 different methods will be tested and evaluated with DCATT during the phases of the DCATT experiment cycle.

3.0 DCATT Optical Systems Modeling

3.1 Optical Schematic and Layout

A simplified optical schematic of DCATT is shown in Figure 3.1. Identified are only those surfaces relevant to an aberrated pupil analysis for single plane diffraction, wavefront sensing and optical control. The following assumptions have been made with respect to the systems modeling in this report:

(1) Static and thermal stress contribute to rigid body misalignments as well as distortion of the shape of the surfaces. Any time constants due to static and thermal stress are thus slow with respect to detector integration times. Alternatively, the control loop bandwidth is much greater than system changes, except for jitter.

(2) The primary effect of jitter is a time dependent tilt of the wavefront yielding rigid body motions of the PSF in the focal plane, note, this assumption has yet to be verified since jitter (dynamics) can induce higher order time dependent aberrations such as time dependent focus, segment to segment piston and tip/tilt errors etc...

(3) The diamond turned and surface micro-roughness contribute to mid- and high-spatial frequencies contribute to a non-deterministic scatter halo which can be incoherently added to the PSF. Strictly speaking this is not true, since, once the surface is manufactured it is totally deterministic. Note, high spatial frequency surface metrology data will be available at sampled locations, thus, statistical information will be available, such as power spectral densities, and probability distributions and can, at a future date, be incorporated into the models. A more deterministic model of the surface scatter would require very high density sampling of the surfaces, however this may be possible.
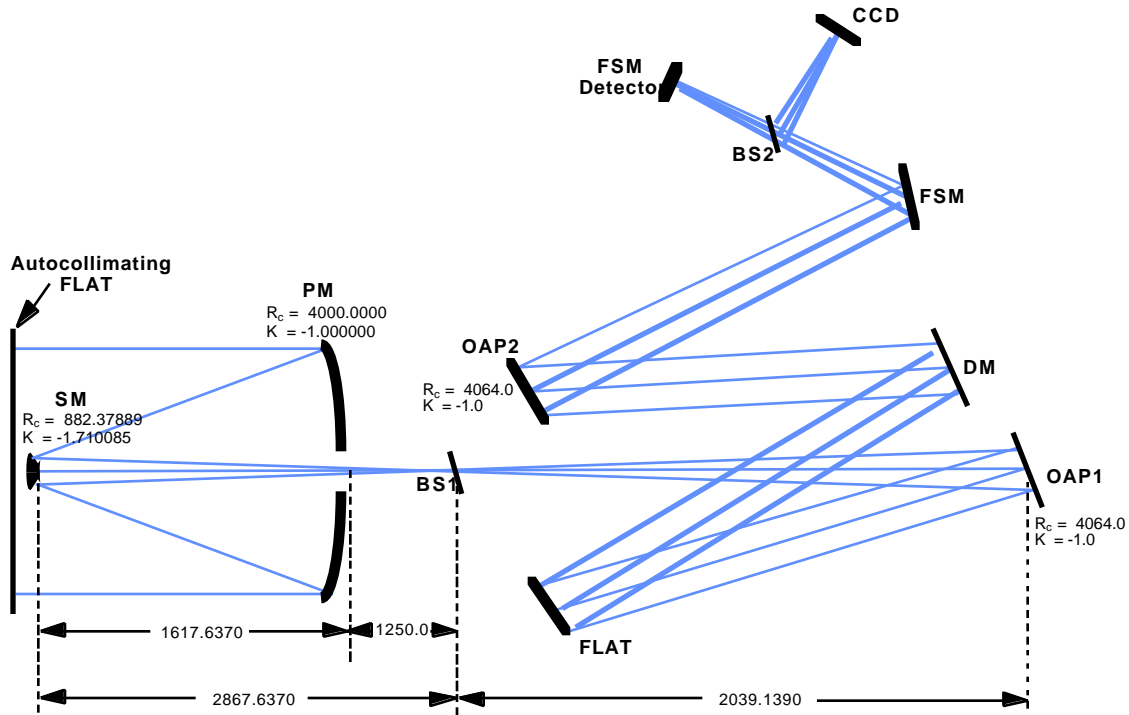
Figure 3.1
Simplified Optical Schematic of DCATT

Not to Scale !
TSM units are mm.

DCATT consists of a: ~1 meter telescope, with a 7 element segmented primary mirror and a monolithic secondary mirror. Each PM segment is a regular hexagon. A focused beam is injected from the source module (not shown in Figure 3.1) at beamsplitter BS1. The source module will initially consist of single laser (=0.6328 μm), and, a broadband Xenon light source with a series of spectral bandpass filters. Future versions of the source module will also have 2 more lasers at =0.543 μm and =0.612 μm The source module provides a diffraction limited wavefront at =0.6328 μm, but cannot provide an aberrated beam to the telescope. The telescope simulator module (TSM) can replace the source module, it has an internal deformable mirror (DM) which can be used to generate an aberrated beam to the telescope. The source module provides an image at the telescope focus of the Optical Telescope Assembly (OTA), which consists of a parabolic primary mirror (PM) and a hyperbolic secondary mirror (SM). The beam is double passed through the OTA via autocollimating off the 1 meter autocollimating flat (AC-flat). The off-axis parabola (OAP1), subsequently collimates the beam which is then reflected by the flat onto the deformable mirror (DM). The DM is at a pupil plane; however, the beam is not

8

at normal incidence on the DM. OAP2 focuses the beam into an f/16.6 cone onto the wavefront sensing camera after reflecting off the fast steering mirror (FSM). Note that initially the FSM will be a fold flat. The optical prescription for the OTA and active optical bench, used in this report, is shown in Tables 3.1 and 3.2 for completeness.

The software used to model DCATT, in this report, is called Lyon's Electro-Optical modeling software, LEO for short. It consists of a series of software modules, some, written entirely in "C", some in MPL, a superset of "C" designed for a MasPar MP2 massively parallel computer. Some of the software has been converted to "C" with Message Passing Interface (MPI), to facilitate portability between both serial and parallel computer architectures. LEO intially began development as a modeling package for the Hubble Space Telescope (HST) in 1988 [10] and expanded [11][12][13] to handle high density raytracing, many Zernike polynomials, polar-Fourier polynomials, detector effects, and jitter following Hubble launch in April of 1990. LEO was further expanded for very high fidelity optical systems modeling and phase retrieval for image deconvolution studies in 1995 [14]. Further refinements including, multiple plane diffraction, rigourous angular spectrum, extended scene modeling, coronagraphs, Shack-Hartmann sensors, segmented apertures, deformable mirrors and detector effects and a multitude of other effects have been added since. Sample input LEO aperture and diffraction files for DCATT are shown in Appendix A for completeness.

The DCATT telescope design is a classical cassegrain with a parabolic primary (PM) and hyperbolic secondary mirror (SM) and as such has no design spherical aberration, however, it does have linear field dependent coma, quadratic field dependent astigmatism, field curvature and distortion. The active optical bench is designed to introduce no new aberrations into the system, but to not passively correct the residual aberrations of the design telescope.

3.2 Wavefront Sensing and Optical Control Considerations

The accuracy, precision, dynamic range, efficiency and robustness of wavefront sensing and optical control for fine figure control loop is affected by:

1. polychromatism
2. jitter, turbulence
3. sampling/pixelization
4. detector MTF
5. scatter, stray light
6. mid- and high-spatial frequencies
7. diamond turning
8. detector dynamic range and quantization
9. noise - readnoise, photon noise, dark current noise, amplifier noise
10. field dependent aberrations
11. phase unwrapping
12. type of method or algorithm
13. calibration
14. wavefront aliasing
15. PM/SM/DM actuator sampling and gains
16. actuator influence functions

In order to determine the effect of each of the above on the wavefront sensing and optical control, a math and computer model has been built and integrated into the LEO package. Each of the above considerations will be developed in more detail in sections 4.0 - sections 5.0. Section 4.0 shows the results and a discussion of the following set of studies:

1. control precision vs. rms WFE
2. control vs. jitter
3. control vs. SNR
4. influence function errors
5. phase diversity ("grid of stars")
6. Shack-Hartmann sensor

In addition the following set of studies are in the process of being performed and these results will be available at a later date:

1. Optimum focus and number of PSFs versus noise decorrelation
2. diamond turning scatter
3. sensitivities (dR, SA3, seg2seg etc...)
4. modal decomposition for optimal control loop calibration
5. float, clamp and slaved actuators (constrained optimization)
6. control metrics, RMS WFE, encircled energy, PSD slope
7. system identification/calibration
8. Extended sources and off-axis sources

Purpose of Study

What is the "best" Wavefront Sensing
and Optical Control Method
for DCATT

•**Determine Effects of:**
- **Calibration**
- **Sampling, finite pixel size,**
- **jitter, noise, diamond turning**
- **Spectral bandpass, Defocus range,**
- **detector MTF, scatter etc...**

- **Comparative Wavefront Sensing**
  - **Compare Phase Ret/Div,**
  - **Shack-Hartmann,**
  - **Interferometry**

•**Parameterize effects with respect to:**
- **Dynamic Range and Resolution**
- **Accuracy/Precision**
- **Sensitivity**
- **Bandwidth of correction**
- **Computational complexity**
- **Robustness**
- **Spatial Frequency Content**
- **Convergence (Multiple start pts.)**
- **Stability**
- **Phase Retrieval/Diversity Algorithms**
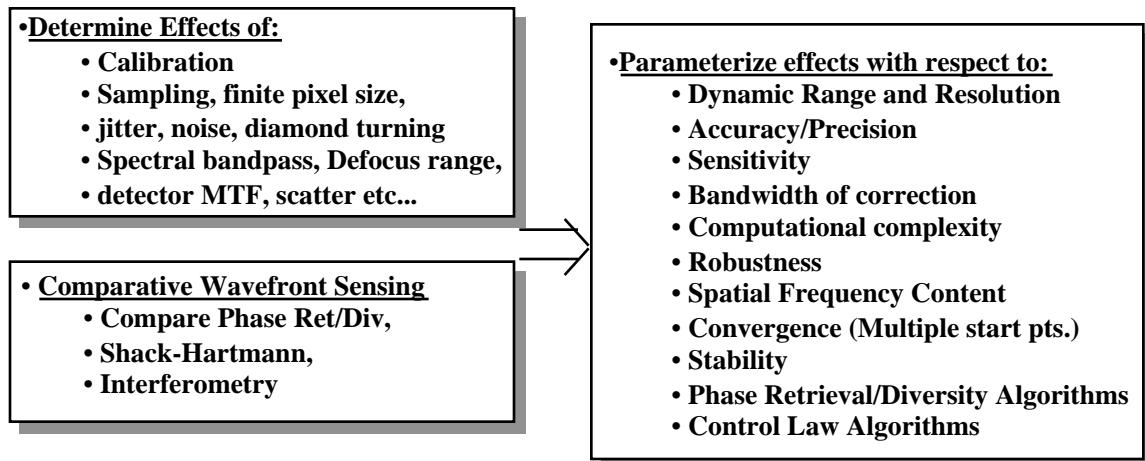- **Control Law Algorithms**

Figure 3.2

DCATT Wavefront Sensing Study

Figure 3.2 shows a schematic of how the ideal study should unfold. We desire to determine the effects of calibration, sampling, pixelization, jitter, noise, diamond turning, power law surfaces, spectral bandpass, scatter, straylight and detector MTF for a number of phase retrieval/phase diversity algorithms, Shack-Hartmann sensing and interferometric methods. We desire to parameterize the effects with regards to accuracy, precision, dynamic range, sensitivity, spatial frequency content and bandwidth of correction. This study

is one step in this process and is not meant to be definitive. More work still needs to be done.

In the following set of studies we utilize the following systems parameters as related to DCATT:

1. Full-Aperture Zernikes /Sub-Aperture Zernikes, {0 - 2.5 rms}.
2. 90.3 cm aperture, 21% obscured, 7 hexagonal segments.
3. Residual Polish Marks / Random power law surfs.
4. Polychromatic PSFs, { =0.6328 um, =10nm}.
5. System Jitter, {0 - 1.5 pixels rms}.
6. Pixelization {9 um pixels}.
7. read noise {13 electrons rms}, photon noise.
8. full-well {80,000 e}, quantization {12 bit = 4096 levels}.
9. Actuator Influence functions (Xinetics 349 DM).

The DCATT optical design was built into the optical modeling software package and the system raytraced for various field angles and foci also as a function of misalignments. Diffraction calculations were also performed. The prescription is shown in Tables 3.1 and 3.2. Appendix A shows input LEO diffraction and segmented aperture files.

nray=	65536	# Number of Rays
nsurf=	11	# Number of Surfaces
fphi=	0	# Polar Field Angle (wrt z-axis)
ftheta=	0	# Aximuthal Field Angle (wrt x-axis)
aper=	903	# System Clear Aperture
irin=	193.09752	# System Inner Aperture (obscuration)
xcent=	0	# X-coordinate of Aperture Center
ycent=	0	# Y-coordinate of Aperture Center
obj=	INF	# Object distance from first surface
lambda=	6.33E-07	# System Wavelength
fill=	0.9203421	# Grid Fill factor (wrt Nyquist)
rtrace=	F	# (T=Tangential, S=Sagittal, F=full 3D)
copd=	Y	# Output ASCII OPD and SPOT ? (Y/N)
bopd=	Y	# Output Binary OPD data ? (Y/N)
cray=	N	# Output ASCII ray state vectors ? (Y/N)
ptilt=	Y	# Remove piston/tilt ? (Y/N)
zfit=	37	# Number of Zernikes to fit.
Segmt=	aper_DCATT.in	# Input segmented aperture file.
Diffn=	pmod_DCATT.in	# Input Diffraction control file.

Table 3.1- LEO System Input File

Surf #	Vertex Radius (mm)	Spacing (mm)	Clear AP (mm)	Conic Constant	Surface Type	Rotation (Degrees)	X-Shift (mm)	Y-Shift (mm)
1	4000.00000	1986.6000	903.0	-1.000000	PARAB	0	0.0	0.0
2	882.37889	-1617.6370	174.0	-1.710085	HYPER	0	0.0	0.0
3	INF	2867.6370	2000.0	0.000000	FOCUS	0	0.0	0.0
4	INF	0.0000	1000.0	0.000000	AXIS-TLT	6.416503	0.0	0.0
5	INF	2032.0000	1064.0	0.000000	DUMMY	0	-227.8	0.0
6	4064.00000	0.0000	164.0	-1.000000	OAP	0	0.0	0.0
7	INF	-1575.0000	155.0	0.000000	FFLAT	5	0.0	0.0
8	INF	1712.5180	137.0	0.000000	DM	-7	0.0	0.0
9	INF	0.0000	1000.0	0.000000	AXIS-TLT	-24	0.0	0.0
10	-4064.00000	-1525.0000	5000.0	-1.000000	OAP	0	0.0	0.0
11	2.00E+10	2032.0277	5000.0	0.000000	FOCUS	0	0.0	0.0

Table 3.2- LEO Input Surface File

3.3 Diffractive Point Spread Function Model

The monochromatic point spread function (PSF) for a single plane aberrated pupil analysis is given by [15]:

$$PSF(x,y;\alpha,\lambda) = \left| \frac{1}{\lambda F} \int A(u,v) \; e^{i\phi(u,v;\alpha)} \; e^{-i2\pi(f_x u + f_y v)} \; du \; dv \right|^2 \quad (3.4.1)$$

$$f_x = \frac{x}{\lambda F}, f_y = \frac{y}{\lambda F} \quad \text{and} \quad \phi(u,v;\vec{\alpha}) = \frac{2\pi}{\lambda} W(u,v;\vec{\alpha}) \quad (3.4.2)$$

where λ is the wavelength; F the system focal length; ϕ the phase due to residual design aberrations, misalignments, surface deformations and residual polish marks etc... Note, the phase can be parameterized in terms of sets of Zernike polynomials, $\{\alpha\}_p$ where each segment can have it's own set. A(u,v) is the aperture mask function, or amplitude of the pupil function, 1 where light is passed, 0 otherwise. Note that A(u,v) will be measured for DCATT and thus once known can be inserted either as raw data or as a functional fit to the raw data.

The intensity integrated across the spectral response and convolved with the detector spatial response is measured in the focal plane. This is commonly referred to as the point response function (PRF).

$$PRF(x, y; \vec{\alpha}) = \int PSF(x, y; \vec{\alpha}, \lambda)\, S(\lambda)\, d\lambda \;**\; rect(\frac{x}{\Delta x})rect(\frac{y}{\Delta y}) \qquad (3.4.3)$$

Where $S(\lambda)$ represents the source spectral radiance multiplied by the spectral transmission of the passband; the detector spatial response function, $rect(\frac{x}{\Delta x})\, rect(\frac{y}{\Delta y})$ is assumed wavelength independent across the filter's narrow spectral passband; Δx and Δy are the x and y pixel spatial dimensions, respectively; and the notation "**" represents convolution. If the detector Modulation Transfer Function (MTF) is known, it can inserted in place of the "rect" functions.

The measurement model is given by:

$$M(x, y; \phi, flux, A, B, C) = flux * PRF(x, y; \phi) + A * x + B * y + C + \eta(x, y) \qquad (3.4.4)$$

flux, represents the observed PRF's integrated flux; A and B represent gradients in the straylight background; C is a constant additive background; $\eta(x,y)$ represents a potentially non-stationary noise component. M is a function of 4 linear parameters (flux, A, B, C) and a non-linear function of the phase. Note that $\eta(x,y)$ is a random variable and, hence, M is also.

Equation (3.4.4) still represents a continuous model for the focal plane image. In order to sample the function we must multiply by a "comb" function, thus, equation (3.4) becomes:

$$M(i\Delta x, j\Delta y; \phi, flux, A, B, C) = flux * PRF(i\Delta x, j\Delta y; \phi) + A * i\Delta x + B * \Delta y + C + \eta(i\Delta x, j\Delta y)$$
$$(3.4.5)$$

Equation (3.4.5) assumes that the stray light and scatter is representable in terms of the 3 background parameters given by {A,B,C}. This assumption will be addressed in more detail in section 4.1.3.5

14

4.0 Wavefront Sensing Studies

4.1 Phase Retrieval

4.1.1 Introduction

In phase retrieval techniques one desires to determine the optical phase delay (i.e. the wavefront error) introduced by DCATT from one or more defocused images. The techniques are well founded theoretically but have not yet seen much use in practice. The Hubble Space Telescope being an exception. The techniques are based upon knowing the optical propagator (Green's function) from an optical pupil plane to the focal plane and it's inverse, see equations 3.4.1 and 3.4.3. In general for an imaging system these are both known and become the Fourier transform kernel [15]. The idea, being to observe or to infer the electric field amplitude in both the pupil and focal planes, and, to estimate the phase in both planes via any number of possible techniques to be discussed below. First shown will be the Iterative Transform Algorithm (ITA) [16]. This algorithm is not used much in practice but is the basic building block of the other phase retrieval and phase diversity techniques discussed here. Thus is behooves the reader to understand it. We also discuss some modifications to this technique and some of the problems with it. We also discuss a focus diversity algorithm (Misell algorithm) [17], a wavelength diversity phase retrieval algorithm, and an extended source phase diversity algorithm (PDA) [18]. We also show the results, and discuss in detail, a series of studies for the focus diversity approach.
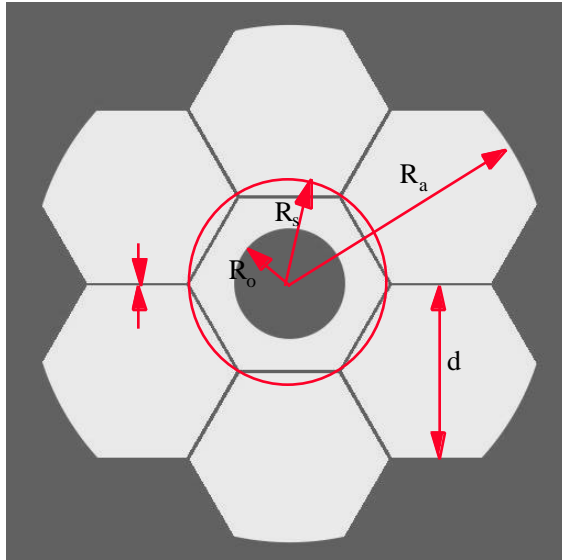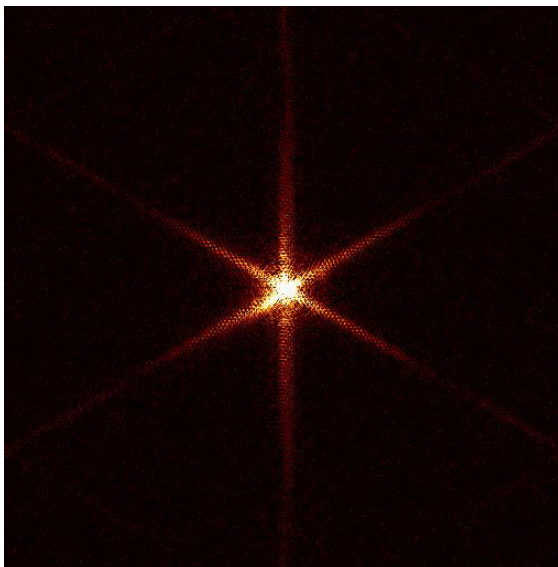
DCATT Aperture Function Geometry
d = 300.000 mm., = 6.000 mm.
R_a= 451.690 mm., R_s= 173.205 mm.
R_o= 96.589 mm.
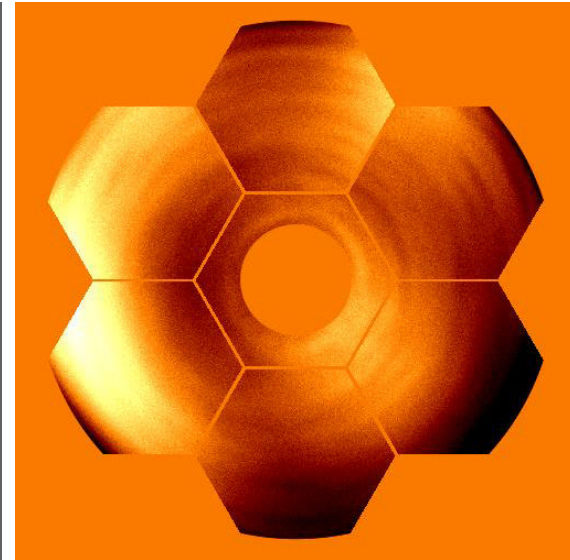
Sample DCATT Phase Function
= 0.285
max = 0.847
min = -1.327

Figure 4.1
Simulated Pupil and Phase Function

DCATT Simulated Sample PSF

Diffraction only
micro-roughness = Y
polish marks = Y
scatter = N
jitter = N
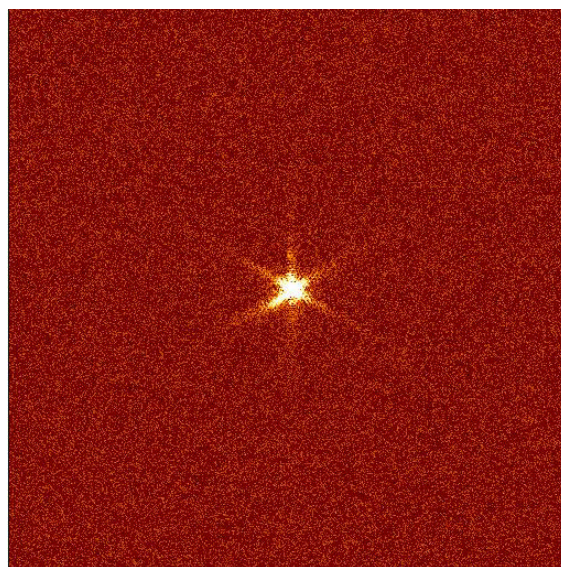
DCATT Simulated Sample PRF

12 bit = 4096
fullwell = 80,000 e
readnoise = 13 e rms
shot noise = Y
detectors = 9 μm

Figure 4.2
Simulated Point Spread Function
and Point Response Function

4.1.2 Modified Iterative Transform Algorithm

The iterative transform algorithm (ITA) [16] is non-parametric method of phase retrieval which treats each point in the phase front as an independent parameter.   In this work we use a modified iterative transform algorithm (MITA). The modification is a method of masking corrupted, noisy and/or saturated data and is based on a method by Fienup [19] and was used by Lyon et.al [14] on HST. In the wings, i.e., far from the core, of an observed PSF, the noise begins to dominate as one moves radially outward from the core. This effect is visible in Figure 4.2. The left side of Figure 4.2 shows a noise free simulation, while the rightside shows the same image with, quantization, readnoise, shot noise and detector MTF.   The detailed structure along the diffraction flares drops into the noise floor. The MITA algorithm attempts to overcome this in regions where the noise dominates. For regions where the noise dominates, when Fourier transforming from pupil space to the focal plane the amplitude of the Fourier transform is retained instead of replacing it with the observed noisy amplitude.   If partial knowledge of the data is available, then one possibility it may be to replace the amplitude with a convex sum of the form:

$$\beta\sqrt{PSF(f)} + (1-\beta)\left|\tilde{P}_n(f)\right| \qquad\qquad (4.1.2.1)$$

where β is the "convexity" parameter such that $0 \quad \beta \quad 1$. In regions of high signal to noise, $\beta = 1$.   In regions of low signal to noise, or regions containing bad data, or near the saturated core, $\beta = 0$. In all cases we used only $\beta = 1$ for valid data and $\beta = 0$ missing or saturated data.
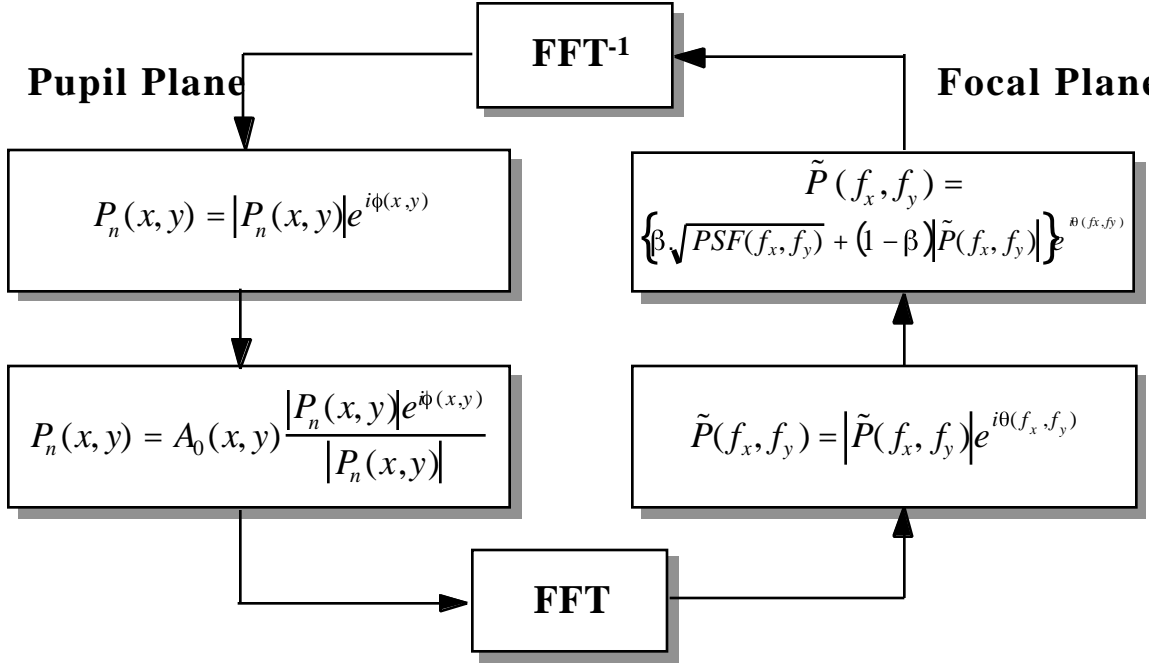
FFT⁻¹

Pupil Plane

Focal Plane

$$P_n(x,y) = \left|P_n(x,y)\right|e^{i\phi(x,y)}$$

$$\tilde{P}(f_x,f_y) = \left\{\beta\sqrt{PSF(f_x,f_y)} + (1-\beta)\left|\tilde{P}(f_x,f_y)\right|\right\}e^{i\theta(fx,fy)}$$

$$P_n(x,y) = A_0(x,y)\frac{\left|P_n(x,y)\right|e^{i\phi(x,y)}}{\left|P_n(x,y)\right|}$$

$$\tilde{P}(f_x,f_y) = \left|\tilde{P}(f_x,f_y)\right|e^{i\theta(f_x,f_y)}$$

FFT

Figure 4.3
Modified Iterative Transform
Algorithm

Figure 4.3 is a pictorial representation of the MITA. The MITA defines a forward and a back propagator from the pupil to focal plane for the complex optical field and assumes an initial starting phase. The forward and back propagators are the Fourier transform kernel and its inverse implemented via FFT techniques. The complex pupil function is Fourier transformed to yield an estimate of the complex focal plane optical field. The focal plane amplitude is replaced with the convex sum of the observed amplitude, $\left(\sqrt{PSF(f)}\right)$, and the amplitude for the Fourier transform, $\left(\left|\tilde{P}_n(f)\right|\right)$, and the phase , $\left(\theta(f)\right)$, is retained. The result is inverse transformed to yield an estimate of the complex pupil plane optical field. The pupil amplitude is replaced with the convex sum of known aperture mask, A_0, and the amplitude of the inverse Fourier transform, $\left(\left|P_n(x)\right|\right)$, and again the phase is retained. This process is cyclically iterated until the phase stabilizes. If near the solution, the MITA will stabilize relatively quickly, otherwise it may take many thousands of iterations. Note if $\beta = 1$ the MITA reduces to the original ITA.

4.1.3 Focus and Wavelength Diversity Algorithms

In general the ITA approach is not utilized in practice since it is not well constrained. In general for an infocus image the problem is ill-posed since the number of free parameters in the phase exceeds the number of valid intensity values in the input PSF, where the valid data points refers to points such that the local SNR is greater than 1. Since the problem is ill-posed any noise in an input variable, must map into multiple phase points, thus correlated noise cells, i.e. regions, are introduced into the phase map. One way to counter this problem is to deliberately introduce a known aberration into the optical path, i.e., a diversity function. This spreads the PSF in the focal plane, and, by integrating longer with the detector we can get many more points with SNR greater than 1. When the number of data input data points exceeds the number of free parameters in the phase, the size of the noise cell drops to ~1 sample point in the phase, thus, the phase noise is decorrelated. Also, by utilizing multiple input PSFs the problem can be made well-posed. The focus diversity (Misell) [17] algorithm, currently baselined for DCATT, uses multiple PSFs, each at different focal positions. In practice we found that with 2 input PSFs, one on each side of focus, it took between 100 to 1000 iterations of the Misell to get an accurate answer, with a 97% convergence rate, i.e., 97% converged with the first starting point for phase. With 4 input PSFs, 2 on each side of focus, we found it took between 10 to 250 iterations with 100% convergence rate.
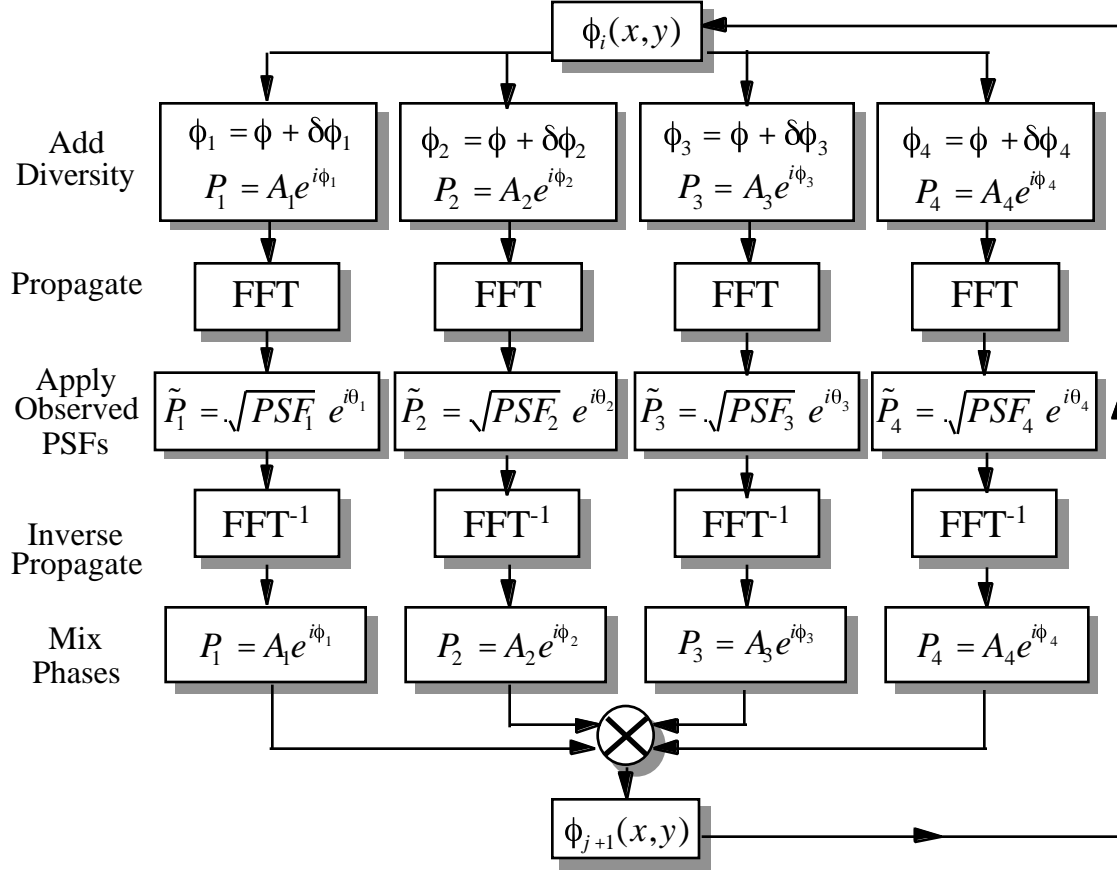
Figure 4.4
4 Point Misell Phase Retrieval Algorithm

Figure 4.4 depicts the 4 point focus diversity (Misell) phase retrieval algorithm used in this study. An initial starting phase is assumed, generated by a uniformly distributed gaussian random number generator on the range [- ,+]. A number of variations are available for the starting phase, (i) we can use a single, random starting phase and subsequently add the diversities, or, (ii) we can use 4 different random realizations of phase, or, (iii) use the phase from the Fourier transform of the square of each of the PSFs. In practice we found that this made little difference to the rate of convergence or to the accuracy of the solution. With regards to Figure 4.4 each of the respective phase functions are inserted into the corresponding pupil functions and FFT'd. This is equivalent to propagating the electric field to the focal plane. The simulated amplitude of the electric fields are replaced with square root of the PSFs for each of the 4 channels and the focal plane phase is retained. The square root of the PSFs represent the observed electric field amplitude in the focal plane.

Each of the 4 focal plane electric fields are back propagated to the pupil plane via the use of an inverse FFT. The complex pupil functions are subsequently mixed to recover the exit pupil phase modulo 2π.
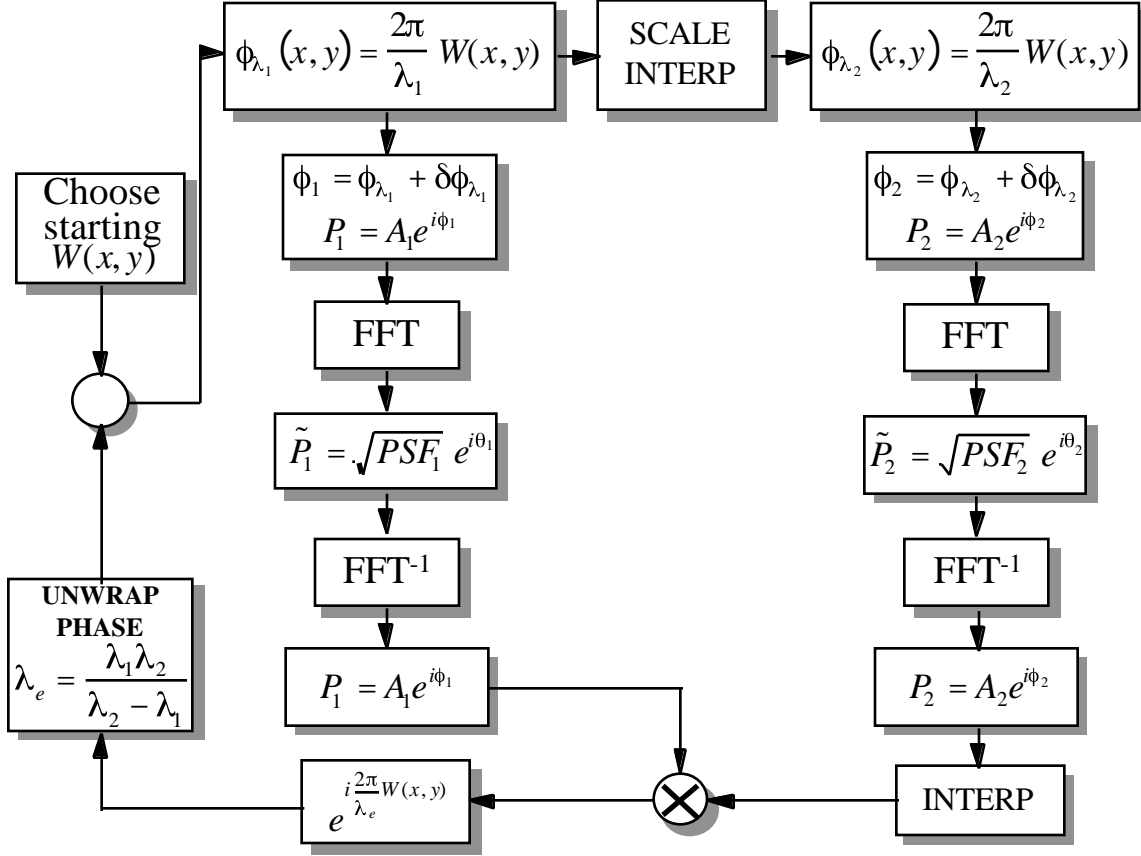

Figure 4.5

2 Point Wavelength Diverse Misell Algorithm

The Misell is a special case of the Phase Diversity [20] approach where the diversity function is focus and the source is unresolved. However, other diversity functions are possible such as shifting wavelength, field position, coded aperture masks, and by deliberately introducing aberrations other than focus. The advantage of using wavelength diversity is resolution of the segment to segment piston problem, and an increased WFE dynamic range as described below.

A two point wavelength diversity algorithm is shown in Figure 4.5. In this approach 2, or more, PSFs are collected each at a different foci and through

different narrowband filters. A starting wavefront, W(x,y), is chosen and the phase generated at each wavelength.   This requires both scaling and interpolation of the wavefront. The interpolation is due to the FFT frequency variable scaling with wavelength, thus, the input pupil function must be sampled on a different grid at every wavelength. The diversity functions are added to the phase. In this case focus was used, thus a parabolic term is added to the wavefront. The complex pupil functions are constructed and propagated to the focal plane. The focal plane amplitudes are replaced with the square root of the observed PSFs, i.e., the electric field amplitude. This updated focal plane field is back-propagated to the pupil via the inverse FFT for each wavelength.   The aperture masks are then applied.   The complex pupil functions are both interpolated to the same grid and subsequently "mixed" together. Mixing refers to taking the complex conjugate of one pupil function and pointwise multiplying times the other one in the following manner:

$$P_1 P_2^* = A_1 e^{i\frac{2\pi}{\lambda_1}W(x,y)} A_2 e^{-i\frac{2\pi}{\lambda_2}W(x,y)} = A_1 A_2 e^{i\frac{2\pi}{\lambda_e}W(x,y)} \tag{4.1.3.1}$$

where the effective wavelength is:

$$\lambda_e = \frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1} \tag{4.1.3.2}$$

Note once the pupil functions are interpolated, $A_1 = A_2$. For λ_1=0.543 microns and λ_2=0.633 microns, the effective wavelength becomes λ_e=3.819 microns, thus, the range over which we do not require phase unwrapping is increased by a factor of ~6, i.e., 3.819/0.633. The effect of multiple wavelength mixing then is an increased dynamic range in phase retrieval, and, resolution of the segment to segment piston problem. Thus we will be able to resolve segment to segment piston differences of up to ~3.8 microns in wavefront error. There is a caveat to this however. If the segment to segment piston difference is larger than 3.8 microns, and, is also an integer multiple of 3.8 microns then it will not be properly resolved. This problem can be circumvented by using more than 2 wavelengths. We believe that the multiple wavelength approach will be more jitter insensitive but have not yet verified this. The effect of jitter on phase retrieval is to introduce unphysical branch points, to be discussed in more

detail in section 4.1.3.2. A branch point at one wavelength will not necessarily be a branch point at another wavelength. Thus, potentially making multiple wavelength phase diversity more robust with regards to jitter. This needs to be investigated in more detail.

4.1.3.1 Phase Retrieval Precision versus Wavefront Error

We conducted a study to determine the phase retrieval precision as a function of the wavefront error (WFE). This gives us a range over which we can expect precise results. LEO was used to generate 100 sets of 4 PSFs, 400 PSFs in all. Each of 4 PSFs in a single set had the same realization of randomly generated full aperture Zernikes, sub aperture Zernikes including piston, tip and tilt. The rms wavefront error was varied over the range 0.0 to 2.5 waves rms WFE. The 4 PSFs in a set each had different foci of -2.0 um, -1.0 um, +1.0 um, +2.0 um rms WFE. The PSFs were polychromatic with a center wavelength of =0.6328 um with a FWHM of =10 nanometers. The detector MTF as well as 512 x 512 sampling grids were folded in. Each of the 100 sets were passed through the 4 point focus diversity algorithm and the phase retrieved. The results were phase unwrapped and subtracted from the known "true" phase and the standard deviation calculated. The standard deviation, in units of waves rms, is plotted versus the input rms wavefront, also in waves rms, in Figure 4.6 The mean standard deviation, i.e., the precision due to the phase retrieval method only, is 0.043 +/- 0.019 microns rms WFE (/23 +/- /52 rms WFE). The anomalous spikes in Figure 4.6 occur due to phase unwrapping and/or segment to segment piston errors. With fine tuning, of the phase unwrapping algorithm, we could remove most of the spikes, however, more work needs to be done to automate this process. Note that the error is relatively constant, after the spikes are removed, over most of the range, with increased precision below 0.25 waves rms input error, also, there is a marked decrease in precision above 2.0 waves rms WFE. This part of the study shows only the precision due to the wavefront sensing algorithm. Precision due to jitter, detector effects, actuators, deformable mirror and control loop will be investigated in the sections following.

The phase from any phase retrieval and interferometric method is recovered modulo 2 if the peak to peak wavefront error is greater than 1 wave. This

"wrapped" phase must be restored to the correct "unwrapped" value. We have developed a method based on graph theory, utilizing a construct known as directed acyclic graph for a very fast, and robust algorithm. We ran a Monte-Carlo simulation result with 100 separate runs of a Misell type phase retrieval algorithm with relatively large excursions of phase error (+/- 8 waves) and found that the algorithm worked successfully in all cases. It works in approximately 0.2 seconds for a 512 x 512 phase map on a single Pentium processor. It also ignores branch points and does not attempt to unwrap them. Figure 4.7 shows the zero to peak WFE versus RMS WFE over the range 0.0 to 2.5 waves RMS WFE for the set of 100 PSFs in this simulation. It is significant to note that the zero to peak wavefront error exceeds 1.0 wave at 0.13 waves RMS. When the zero to peak exceeds 1.0 waves, i.e., the phase range exceeds 2 , phase unwrapping is required, thus it appears likely that for DCATT we will need phase unwrapping. The conversion from RMS to Zero to Peak, as seen from Figure 4.7, is 7.5 waves zero to peak to 1 wave RMS.
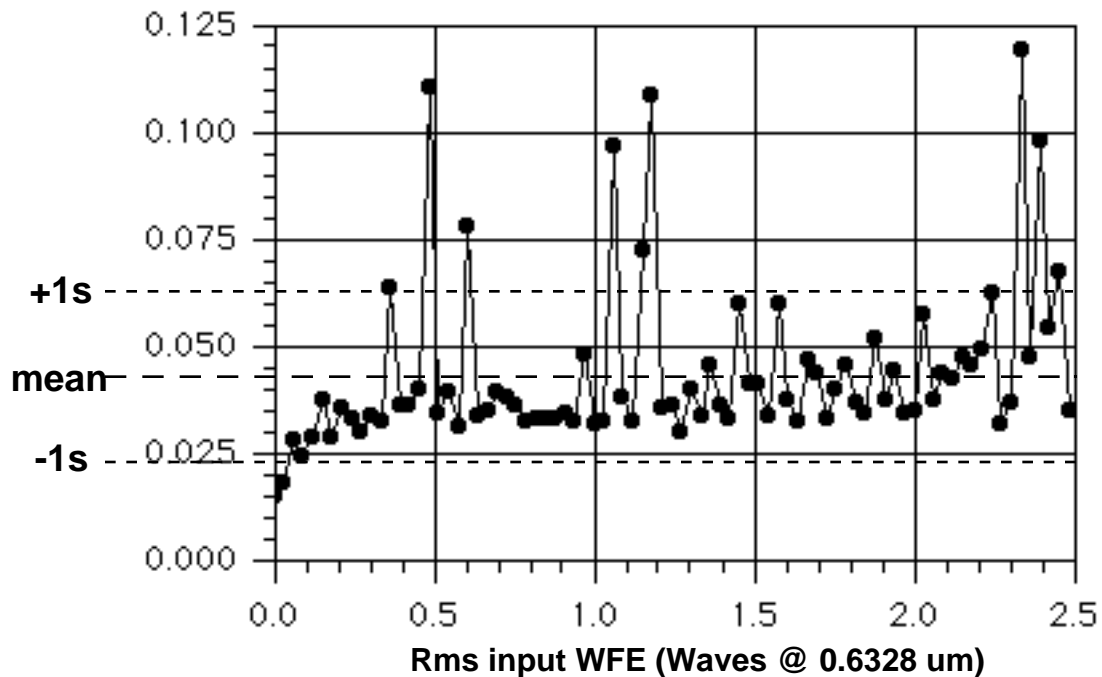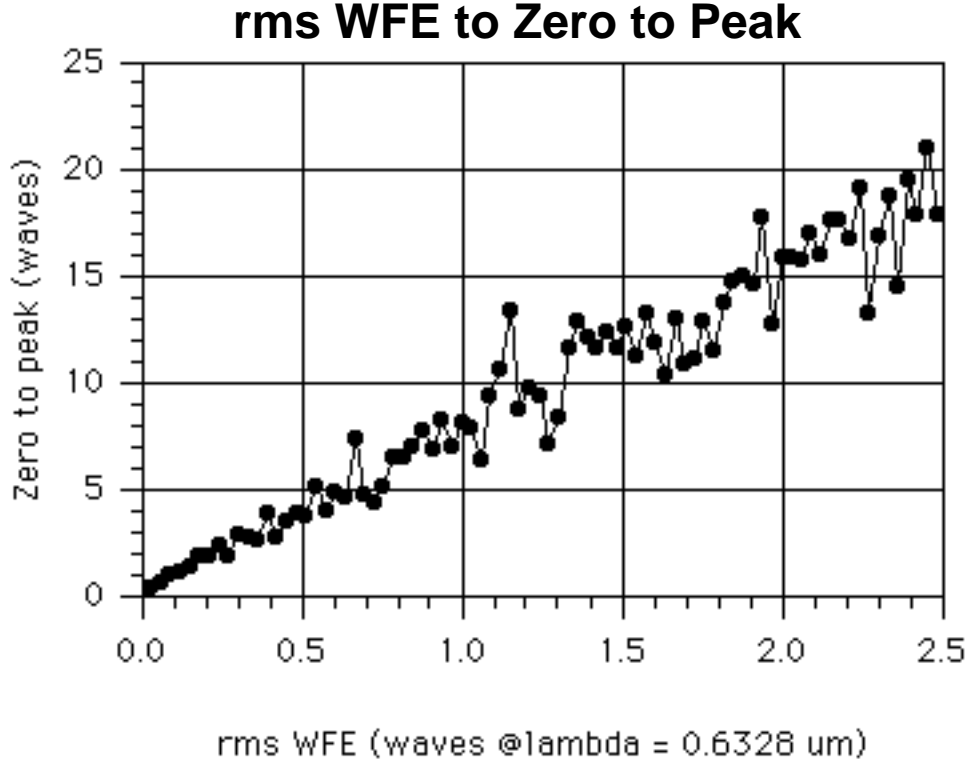

Figure 4.6
WFS Precision vs Input WFE

rms WFE to Zero to Peak

Figure 4.7
Zero to Peak versus RMS Wavefront Error

## 4.1.3.2 Jitter, Phase Unwrapping and Branch Points

In this section we investigate the effect of jitter on the 4 point phase retrieval algorithm. The jitter model currently utilized in this study is a non-deterministic white noise jitter in the transverse directions only. Thus, this assumes only that the jitter shows up as a line of sight jitter and is modeled by a low pass filter on the optical transfer function. The filter function is given by:

$$\left\langle H(f_x, f_y) \right\rangle = e^{-\left(\sigma_x^2 f_x^2 + \sigma_y^2 f_y^2\right)} \qquad (4.1.3.2.1)$$
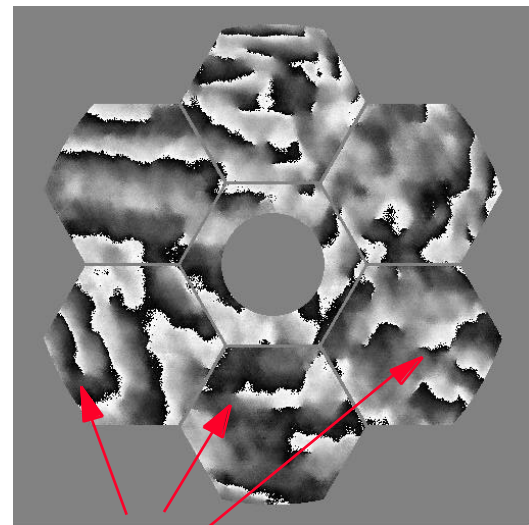
and represents an ensemble averaged jitter, or, alternatively jitter which is relatively fast with respect to the detector integration time. σ_x and σ_y represent the rms jitter in the x and y transverse directions respectively. More work needs to be done in the jitter area. This model needs to be expanded to include actual jitter along the line of sight, potentially, giving time

dependent focus and also, segment to segment jitter in terms of time dependent pistons and tilts. These models will be developed as time permits. Jitter is also being measured at various locations in the lab and it may be possible to fold the observed jitter into the models.

It was found during the simulations that the primary effect of jitter for DCATT was the introduction of unresolvable branch points into the retrieved phase. The left side of Figure 4.8 shows a phase retrieval result with an input of 4 PSFs each at different foci each with 0.1 pixels (9.0 um pixels) rms random jitter and the right side of Figure 4.8 shows the results of phase retrieval also with a 4 input PSFs at the same foci but with 1.0 pixels of rms jitter. The left figure shows well defined regions of phase; the boundaries are 2 jumps prior to phase unwrapping. The right figure contains qualitatively similar regions, however, there are unresolved branch points. A branch point is a point such that if we define a closed contour about it and traversed the contour and summed the 2 phase jumps we would get a 2 phase jump. Note that the contour must not cross the segment boundaries or extend outside a segment. The branch points imply that phase retrieval stagnates at an unphysical solution. Phase unwrapping cannot unwrap branch points since there is no solution at these points.
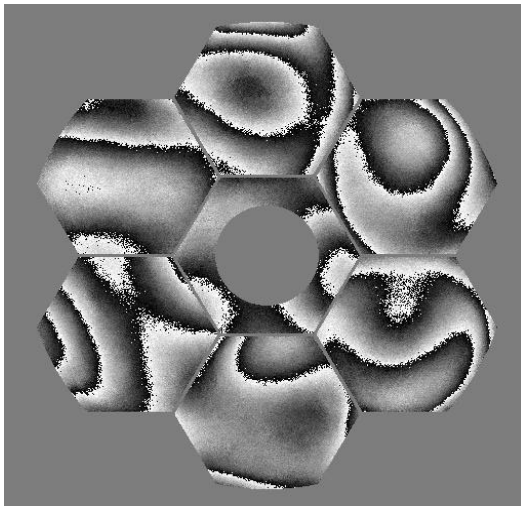
Jitter Induced Branch Points

**WFE = 1.00 waves rms
Jitter = .1 pixel rms (9 um)**

**WFE = 1.00 waves rms
Jitter = 1 pixel rms (9 um)**

**Figure 4.8
Jitter Induced Branch Points**

In order to determine the tolerable amount of jitter, we conducted a study to determine the phase retrieval based wavefront sensing error, for the 4 point focus diversity algorithm, as a function of jitter and input wavefront error. The results are plotted in Figure 4.9. Three sets of 4 PSFs each were generated. The first set of 4 had 0.25 waves RMS wavefront error, the second with 0.50 waves RMS wavefront and the third with 1.00 waves RMS wavefront error. The wavefront error was a random mix of full and sub aperture wavefront errors and the PSFs were all polychromatic with =0.6328 um and FWHM of 10 nanometers. The sets of PSFs were passed through the 4 point focus diversity algorithm and subsequently through phase unwrapping and the retrieved result subtracted from the known "true" phase. The input jitter was varied from 0.1 pixels (9 um pixels) rms to 1.50 pixels rms and the phase retrieval error versus rms jitter tabulated in Figure 4.9. Note that the error increases both with increasing jitter and increasing input WFE. One possible way to make phase retrieval more robust with regards to jitter is the wavelength diversity algorithm described in section 4.1.3. This needs to be investigated in more detail.
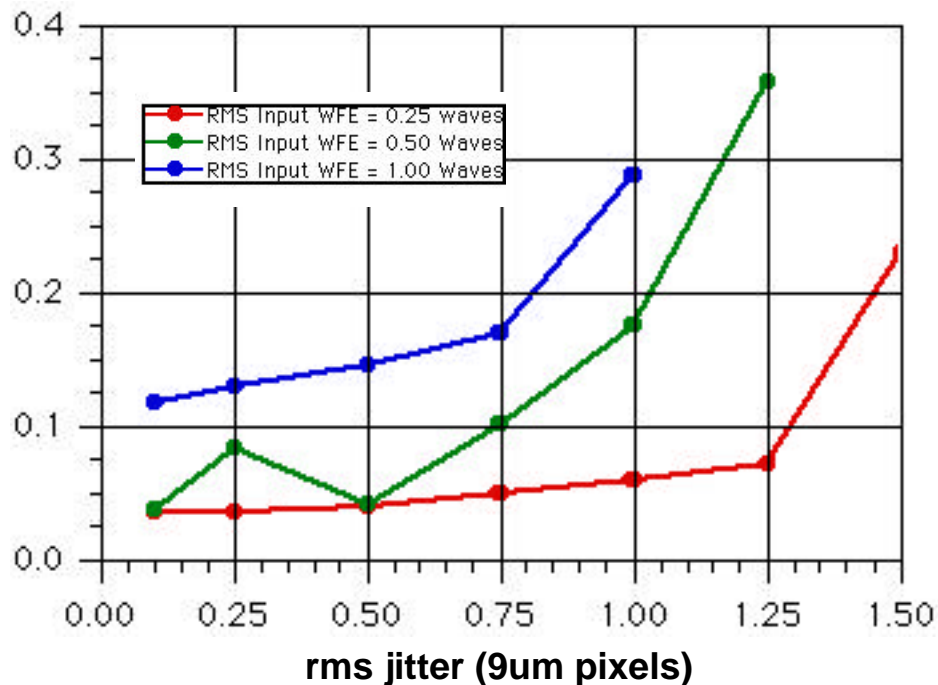

Figure 4.9
Wavefront Error vs. RMS Jitter

4.1.3.3 WFS Error versus Signal to Noise Ratio

In order to determine the effect of noise we conducted a study with the 4 point focus diversity algorithm to determine the error introduced as a function of signal to noise ratio (SNR). Two sets of 4 PSFs were generated. The first set contained 0.25 waves RMS wavefront error and the second set 0.50 waves RMS wavefront error. The 4 PSFs in each set had different focal errors of {-2.0, -1.0, +1.0, +2.0} waves rms. The effects of detector sampling, finite detector element size (pixel MTF), quantization error (12 bit), read noise (13 electrons rms), and shot noise such that the integration time was varied to make the peak electron counts vary from 4000 to 80,000 electrons in the peak pixel. Each of the realizations were passed through the 4 point focus diversity algorithm and the retrieved phase subtracted from the known "true" phase and the error tabulated versus peak electron counts plotted in Figure 4.10. The error was less that 0.10 waves rms for 8,000 electrons peak and seemed to be relatively independent of wavefront error for greater than 8,000 electron counts peak. Note that in this study the integration time was adjusted separately for each of the 4 defocused PSFs, in a given set, such that peak met the graphed peak amount. This may not be the case in the lab, but, it is more ideal since all PSFs would have the same dynamic range.
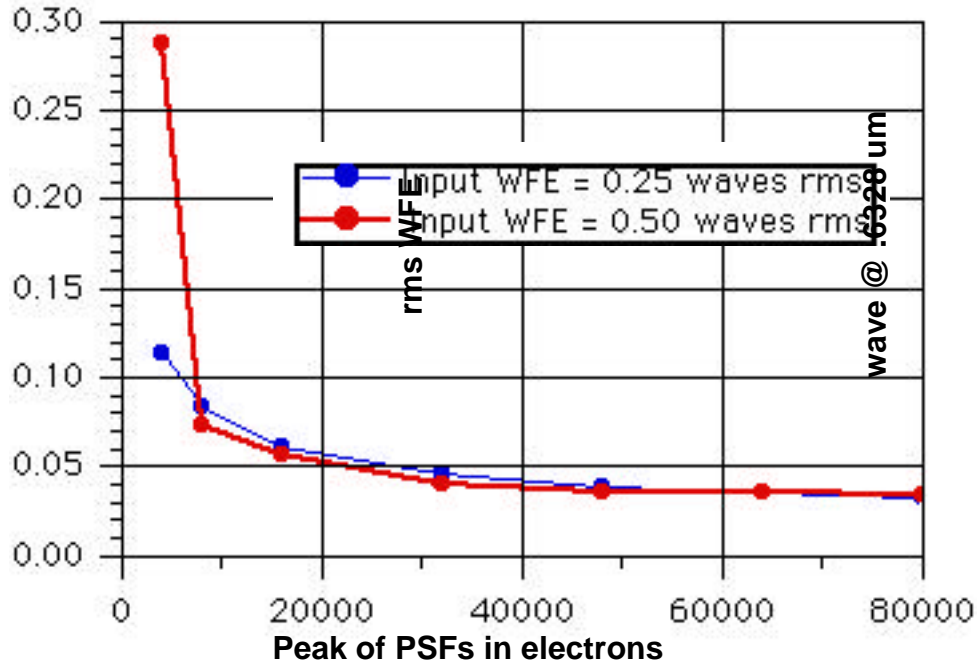
Figure 4.10
Wavefront Sensing Error vs. peak Electron Count

4.1.3.4 Deformable Mirror Influence Function Errors

The deformable mirror influence functions may initially not be accurately known, and subsequently may drift with time. We performed a study to ascertain the sensitivity of WFE to errors in the influence functions. The influence function response was modeled as $R(r) = e^{-ar}\cos(ar)$. Where "r" is the radial variable centered on an actuators location of the deformable mirror grid, "a" is the actuator coefficient, fixed, and assumed known for a given actuator. In this model each actuator is assumed to have the same response, and, the surface is a linear superposition of all the actuator influence functions. We simulated the entire control loop with 0.25 waves of input rms WFE and passed it through the entire control loop simulation. We subsequently varied the value of "a" over the range of {-25%, +25%} from nominal and determined the rms WFE introduced by this process. It was found, surprisingly, that for we could tolerate relatively large errors in "a", with the system being very insensitive over the range {-5%, +5%}. Figure 4.11 shows a plotted of rms WFE versus fractional coefficient error.

29

@ 0.6328 um
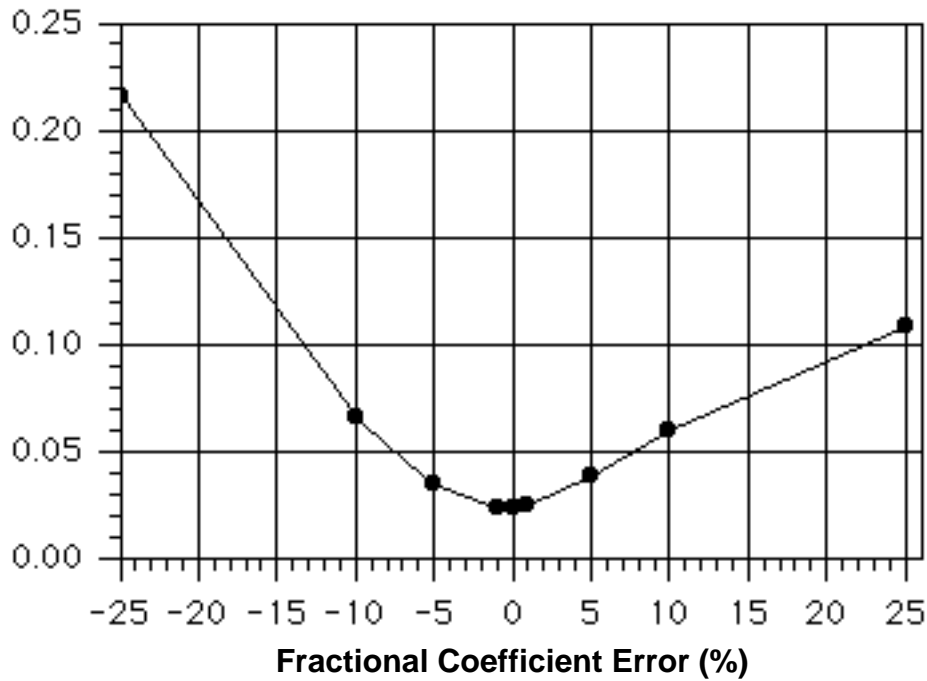
Figure 4.11
Wavefront Error vs. Fractional Influence Function Errors

4.1.3.5 Diamond Turning, Mid- and High-Spatial Frequencies

In this section we develop models for the mid- and high-spatial frequency content, due to polishing, diamond turning and micro-roughness. The detailed math models are developed in Appendix B. As of this writing we have not simulated the effects of diamond turning on the image quality, however, we have simulated mid spatial frequencies and micro-roughness. We have used a random power spectral density law for mid spatial frequencies and micro-roughness. However, at this time it is not known whether these models have the correct parameters until after the DCATT final polishing and metrology process.

The effects of mid-spatial frequencies, due to residual polish marks, were included by using a set of polar-Fourier polynomials of the form:

$$\sum_{n=0}^{N}\sum_{m=0}^{M} p_{n,m}^{(1)} C_n(r)C_m(\theta) + p_{n,m}^{(2)} C_n(r)S_m(\theta) + p_{n,m}^{(3)} S_n(r)C_m(\theta) + p_{n,m}^{(4)} S_n(r)S_m(\theta) \qquad (4.1.3.5.1)$$

where

$$C_n(r) = \cos\frac{2\pi n}{1-\varepsilon}(r-\varepsilon) \quad , \quad S_n(r) = \sin\frac{2\pi n}{1-\varepsilon}(r-\varepsilon) \quad \text{for} \quad n = 0,\ldots, 13 \qquad (4.1.3.5.2)$$

$$C_m(\theta) = \cos(m\theta) \ , \quad S_m(\theta) = \sin(m\theta) \quad \text{for} \quad m = 0,\ldots, 16 \qquad (4.1.3.5.3)$$

where the azimuthal index is m and the radial index is n; $\varepsilon = 0.21$; and $\{p_{n,m}^{(1)}, p_{n,m}^{(2)}, p_{n,m}^{(3)}, p_{n,m}^{(4)}\}$ represents the set of randomly chosen coefficients. When n=m=0, only one term remains in equation (4.1.3.5.1); and when only n or m equals zero, only two terms remain. This gives a total of 4NM+2N+2M+1 terms; when N=13 and M=16 the terms total 891. Note that n ranges from 0 to 13 and thus gives 13 cycles across the radius of the annulus or greater than 26 cycles across the complete aperture. The polynomials do not form an orthogonal set, but in practice, it was found that mid spatial frequency structure could be modeled more easily, i.e., with fewer terms, using the polar-Fourier polynomials than a large number of Zernike polynomials. This is believed to be true since the polishing process residuals contain strong radial and azimuthal periodicity. The randomly chosen coefficients are chosen to follow a radial power law power spectral density. These residual polish marks are evident in the right side of Figure 4.1 as the approximately concentric mid-spatial frequency rings. At this time we assumed a random power spectral density law, however, as the primary and secondary mirror metrology data becomes available it can be folded in to give a more realistic simulation and for subsequent model verification when real images become available from DCATT.

The current post-polish rms mirror requirement for the DCATT primary mirror (PM) is 50 - 80 Angstroms rms surface error over the range of periods from 2 microns to 200 microns. Three possible approaches have been investigated for modeling the effect the diamond turning process on the PSF. The first approach is based on the small angle approximation and on knowledge of the surface statistical properties. The second approach does not use the small angle approximation, but does require the surface statistics to be

normally distributed and the surface power spectral density (PSD) must be known or inferred. Both of these approaches are ensemble averaged approaches, and, as such predict the scatter over an ensemble of DCATT's. The third approach requires high density metrology of the surface, but, does not give an ensemble averaged result. Each of the 3 approaches are mathematically derived in Appendix B. This work is currently ongoing and more results will be forthcoming.

4.1.4 Phase Diversity

4.1.4.1 Introduction

Phase retrieval utilizes as input one or more PSFs, which are actually images of unresolved sources at one or more foci, and/or wavelengths. Phase diversity, on the other hand, allows for an arbitrary input source, albeit through a narrowband spectral filter(s). Phase diversity should be evaluated for use on NGST, possibly, as the primary method wavefront sensing for fine figure control, and, for in-situ monitoring of wavefront quality during science observations. It is possible that the telescope image quality could drift since many NGST observations could be long. Thus, it behooves us to investigate phase diversity for NGST by a comprehensive literature search, computational modeling and simulation of a limited subset of the methods, and, a set of phase diversity experiments on DCATT. Furthermore, the computational modeling aspect should be predictive in nature, i.e., it should attempt to predict the wavefront and controls performance of DCATT. After subsequent DCATT experiments have been performed, the models should be updated to reflect actual system performance. In this manner we gain confidence in our models with the goal of extrapolating them to NGST.

In this section we first describe some of the results of previous phase diversity studies, in the context of their applicability to DCATT. Following this we describe what still needs to be done to adequately characterize phase diversity for DCATT and NGST. We also investigate a two channel phase diversity algorithm which we have coded and have begun some DCATT simulations.

4.1.4.2 Previous Studies

Paxman and Crippen [21] and Paxman and Fienup [22] studied the phase diversity technique for sparse-aperture 6 element phased array telescopes with piston errors, additive Gaussian noise and quantization errors with regards to 18 misalignment parameters. They also discussed the use of a Wiener-Helstrom filter as the inverse filter. If prior knowledge is available about the objects power spectrum then this is a more optimal filter for object estimation.

Kendrick et.al. [23][24] describe a phase diverse wavefront sensor for imaging systems which uses 2 separate CCD array focal planes, which records, simultaneously, the best focus image and a defocused image of the same source. The primary emphasis of this study being to evaluate the phase diverse sensor with real extended scene data, and, attempt to increase the sensor's temporal bandwidth and reduce noise sensitivity. They used a 19 element, hexagonal segmented mirror and also used 7 segments to simulate a segmented primary mirror. The Kendrick et.al. study was primarily to correct for atmospheric turbulence, thus, system bandwidth was a major issue. They utilized a General Regression Neural Network (GRNN) to calculate the wavefront errors, and, subsequently to move the segmented optics to minimize wavefront error. The GRNN requires an extensive training set which must be stored. Also developed were a number of interesting alternative metrics to assess performance and drive the control loop. A series of open and closed loop piston tests were conducted. The tests had random piston errors on the segments over the range [-0.3, +0.3] waves wavefront error at a wavelength of 0.6328 um. The results showed they could recover the input scenes modulation transfer function reasonably well and that the wavefront (i.e., piston only) was recovered to ~0.064 (/16) waves rms for a 4 segment open loop test and ~0.038 (~ /26) closed loop. A 6-segment closed loop test gave errors of ~0.065 (~ /16). This primary reason for the drop in precision with a larger number of segments is that the density of the training set for the GRNN is lower. This study is interesting in that it uses actual hardware and a segmented mirror in a closed loop fashion, however, only piston error were used. We would like to see results for higher order aberrations, and, ideally a power spectral density of errors versus wavefront spatial frequency. One tantalizing result, demonstrated in [23] , suggests that phase diversity techniques developed for monochromatic light will work with white light as well. One advantage of the GRNN is that it can be trained with white light images, thus using all the photons, and it has been demonstrated at bandwidths of up to 10 Hz.

Lee, Roggemann, Welsh and Crosby [25]   evaluate a Gonsalves [20] phase diversity technique for space telescope wavefront sensing via a series of Monte-Carlo simulations.   Interestingly they use a strawman 8 petal NGST model to evaluate their method for a number of different photon limited imaging scenarios, including full aperture Zernike modes, piston and tip-tilt

segment errors. However, they neglected any segment figure errors. Their study addressed the photon limited case (Poisson noise) with an input unresolved point source. The authors chose an ensemble of 100 input wavefronts with errors which were all ~ /10 rms and ran a Monte-Carlo simulation with a 2 channel Gonsalves based phase diversity algorithm with a known focal difference between the channels. There results for the wavefront estimation errors range from a worst case of 0.057 waves (~ /18) for the case of 1000 photons per image to a best case of 0.002 waves (~ /500) for the high photon count case. Note that the authors point out that the Gonsalves approach is a maximum-likelihood estimator for Gaussian noise, not Poisson, and thus this approach could be improved on. This study did not take into account, jitter, quantization error, readnoise, and figure error on individual segments, it primarily addressed the photon limited case.

In Paxman, Seldin et.al. [26][27][28] two alternative phase diversity methods are developed and compared for solar image restoration. The first approach uses a maximum-likelihood estimator for Poisson noise, and, hence, in principle, would be more optimal for the low light case. The second approach introduces a noise filter and preferentially weights the different image channels based on signal to noise. Also, the numerical optimization methods are significantly different. The authors compare the results of a restoration of 100 solar image pairs with both methods. An error analysis is described for both the image restoration and the wavefront sensing error. The image restoration has relative mean difference error of 1.5% +/- 1.0%, i.e. relative rms difference in restored objects between the 2 methods. The rms difference in recovery of the wavefront is 0.043 (/23) waves rms. Note, however, this is not an absolute error since it compares 2 different methods, any, systematic errors which occurred in both methods, may tend to decrease the precision. However, due to the high degree of correlation in both the restoration and wavefront sensing we believe it is a good measure, also, 2 separate teams implemented the 2 different methods potentially mitigating against systematic errors.

Paxman et.al. [29] is a nice theorectical development of phase diversity for both the Poisson and Gaussian noise cases. Discussed also is regularization and optimization methods.

The sequence of studies described above is not meant to be comprehensive, but gives an overview of the current state of phase diversity. It has potential as a wavefront sensing method for DCATT and NGST, particularly with extended sources. It has the advantage that a point source need not be identified and "cut" from the observing scene, whatever is available potentially could be used, making this method a natural for in-situ observations of wavefront quality if some type of known diversity can be introduced into the system.

A number of things are missing from the set of studies performed so far. We need to know the performance of phase diversity as a function of source spatial frequency content. An input point source contains all spatial frequencies, the optical system is a low pass filter, attenuating the high frequencies, representing a loss of information for spatial frequencies such that the SNR is less than 1.0. This loss of information occurs at lower spatial frequencies for extended objects. This implies, intuitively, that the error in wavefront spatial frequency recovery will grow with increasing spatial frequency. This wavefront sensing transfer function needs to be determined and parameterized with respect to input scene spatial frequency content. The previous studies also don't take into account figure errors on a per segment basis, system jitter, and segment to segment jitter, all of which are important to DCATT and NGST.

A series of DCATT experiments could be done by imaging a mask into the first focal plane of DCATT, the mask, e.g., should be at first unresolved and a series of experiments conducted with increasing mask size. Also with a grid of randomly located unresolved sources or a mixture of unresolved and resolved sources, to simulated crowded astronomical fields. The mask illumination must be spatially incoherent.

In theory an arbitrary input source, which is actually unknown, can be used since from a purely mathematical point of view the method becomes source independent. In practice we find the accuracy of wavefront spatial frequency content recovery is strongly dependent on the source's spatial frequency content. Generally, as the source extent becomes large, the spatial frequency content is lower and the accuracy of the wavefront sensing decreases with

increasing spatial frequency. Low spatial frequency wavefront error is recovered rather accurately, however higher spatial frequencies have errors that increase rapidly with spatial frequency.

4.1.4.3 Phase Diversity Simulation

We have developed the phase diversity algorithm detailed in [18] and some simulations to show the decrease in wavefront sensing precision as a function of increasing spatial frequency for a single input scene. Appendix C develops the mathematics of the phase diversity algorithm we used here. Figure 4.12 is a flowchart of this algorithm described here. The method is as follows for a 2 channel phase diversity method with focus as the diversity function. An image is taken with the sensor defocused, alternatively the focal plane can be moved. Another image, of the same object, is taken at another focus position. The absolute focus doesn't have to be known, however, the difference in focus between the 2 channels must be accurately known. Both channels must "see" the same aberrations, and the same object, except for a known difference in focus. Note, that the diversity aberration doesn't have to be focus, it could be another aberration, or field dependence and/or wavelength, or even a coded aperture. Generally, in practice, the more diversity the better the result. With regards to Figure 4.12 the algorithm proceeds as follows, a "reasonable" estimate for the starting phase is chosen. "Reasonable" refers to using a starting point as close to the solution as possible. The closer the starting point the faster it converges. The diversity functions are added to the starting phase for each of the channels and the pupil function constructed and propagated to the focal plane via Fast Fourier Transforms (FFT). An estimate of the object is made in the Fourier domain via a multiple channel Wiener filter. The object is inverse FFT'd and the constraint that the object must greater than or equal to zero is used. Any point in the object less than zero is set to zero. This is known as a projection operation. Then each of the PSFs are estimated via separate single channel Wiener filters and again if any PSF points are less than zero they are set to zero. An iterative transform algorithm is used to propagate from focal plane to pupil plane and the pupil functions are estimated. The 2 (or more) pupil functions are mixed together and an updated estimate of the phase is returned. This is fed back at the top of Figure 4.12 and the process cyclically iterated until both the object and the phase estimates are no longer changing.

$$\phi^{(j)}(x)$$

$$\phi_1 = \phi + \delta\phi_1$$
$$P_1 = A_1 e^{i\phi_1}$$

$$\phi_2 = \phi + \delta\phi_2$$
$$P_2 = A_2 e^{i\phi_2}$$

$$PSF_1 = |FFT(P_1)|^2$$

$$PSF_2 = |FFT(P_2)|^2$$

FFT

FFT

$$O(f) = \frac{P\tilde{S}F_1^*(f)d_1(f) + P\tilde{S}F_2^*(f)d_2(f)}{|P\tilde{S}F_1(f)|^2 + |P\tilde{S}F_2(f)|^2 + \beta(f)}$$

$$O(x) = FFT^{-1}\{O(f)\}$$
$$if \ O(x) < 0 \ then \ O(x) = 0$$

$$P\tilde{S}F_1(f) = \frac{O^*(f)d_1(f)}{|O(f)|^2 + \alpha(f)}$$

$$P\tilde{S}F_2(f) = \frac{O^*(f)d_2(f)}{|O(f)|^2 + \alpha(f)}$$

$$PSF_1(x) = FFT^{-1}\{PSF_1(f)\}$$
$$if \ PSF_1(x) < 0 \ then \ PSF_1(x) = 0$$

$$PSF_2(x) = FFT^{-1}\{PSF_2(f)\}$$
$$if \ PSF_2(x) < 0 \ then \ PSF_2(x) = 0$$

ITA
(see Figure 4.3)

ITA
(see Figure 4.3)

$$P_1 = A_1 e^{i\phi_1}$$

$$P_2 = A_2 e^{i\phi_2}$$

$$\phi^{(j+1)}(x)$$
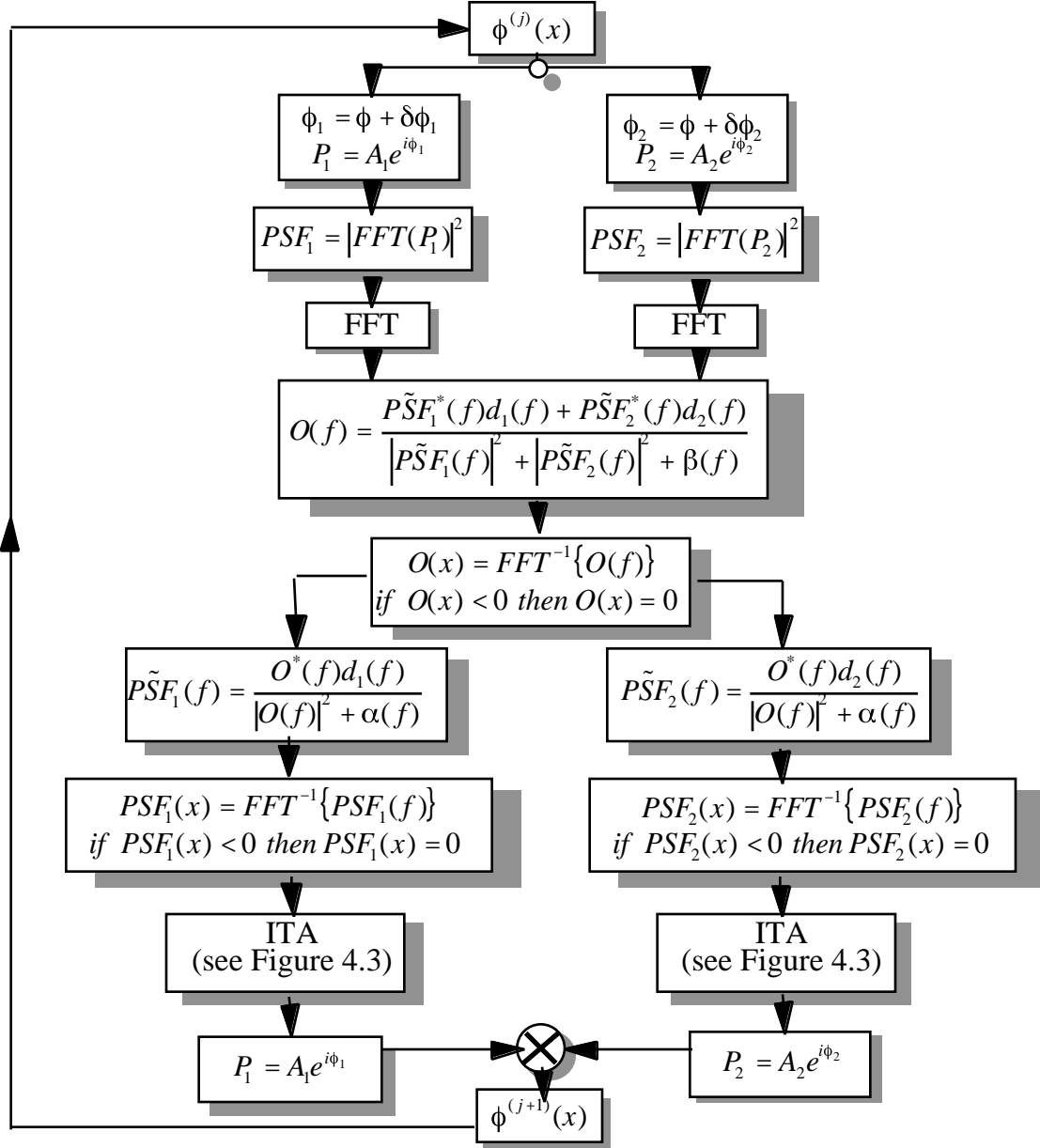
Figure 4.12

2 Channel Extended Source Phase Diversity Algorithm

38

Image1⇒

PSF1
(2X)

PSF2
(2X)

Image2⇒

Grid of Sources
(Object)
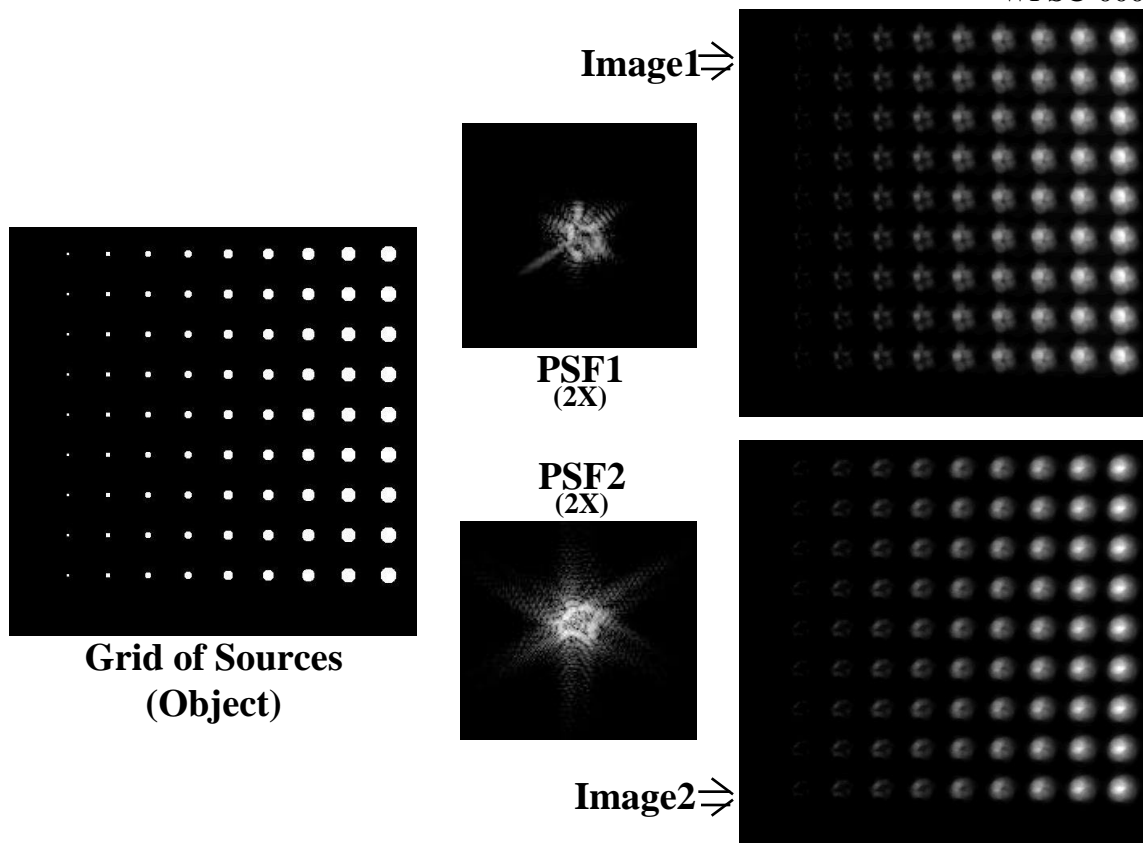
Figure 4.13

Input Images for DCATT 2 Channel Phase Diversity
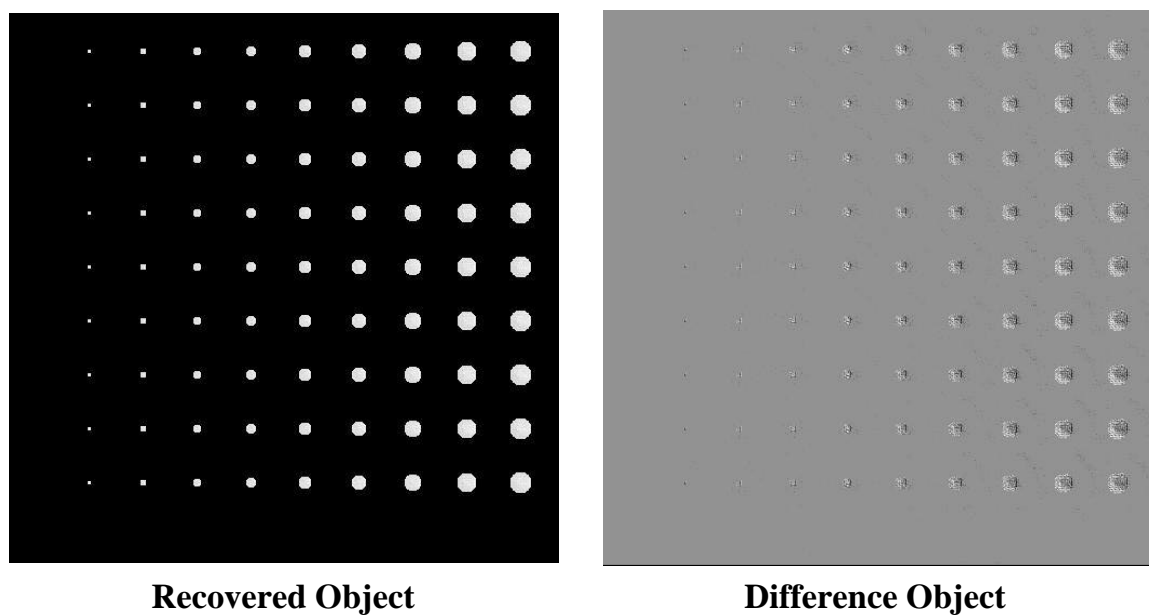
Recovered Object

Difference Object

**Figure 4.14 - Output and Difference Object**

Figures 4.13 and 4.14 shows the results of the 2 channel extended source phase diversity algorithm with a simulated input object. The leftside of Figure 4.13 shows and input grid of sources. The source grid is convolved with PSF1 and PSF2 and the outputs are shown as image1 and image2 of Figure 4.13. The PSFs both have the same realization of aberrations but with a +/- 1.6 waves of defocus between them, i.e. PSF1 is at -1.6 waves RMS WFE of focus and PSF2 is at +1.6 waves RMS WFE of focus. The aberrations have random full and sub-aperture Zernikes. The algorithm shown in Figure 4.12 is iterated approximately 1000 iterations to arrive at the result for the object in the lefthand side of Figure 4.14. The recovered phase is shown in the leftside of Figure 4.15. The rightside of Figure 4.14 shows the difference between the recovered object and the "true" object (leftside of Figure 4.13). The rightside of Figure 4.15 shows the "true" phase and the bottom of Figure 4.15 shows the difference between the recovered and the "true" phase. At first glance this looks like white noise, however, it is missing low frequency content. Figure 4.16 shows a 1 dimensional slice through the power spectral density of the bottom of Figure 4.15. The plot in Figure 4.16 is a log-log (base 10) plot, notice is starts rising at greater than 1.0 cycle per aperture, with the most severe errors occurring at increasing spatial frequencies. Qualitatively from the upper left of Figure 4.15 we see a type of speckle noise introduced into the recovered phase. We believe that with further research we can fine tune this algorithm, and/or come up with a better approach.
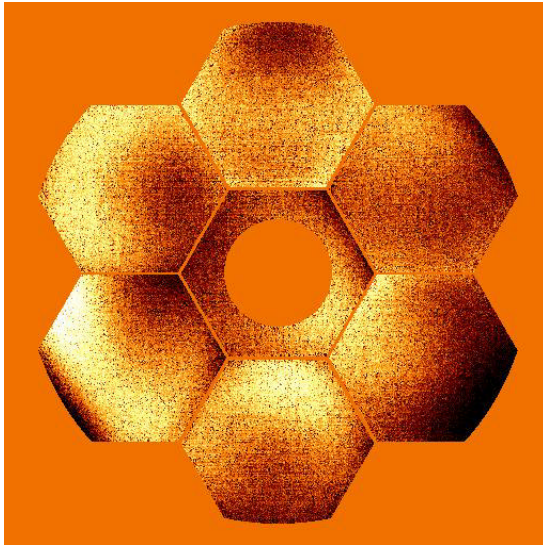
**Phase Diverse
Recovered Phase
0.320 λ rms**
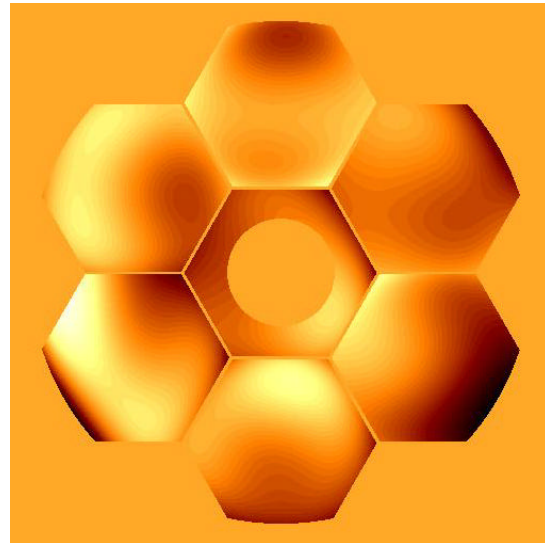
**True Phase
0.273 λ rms**
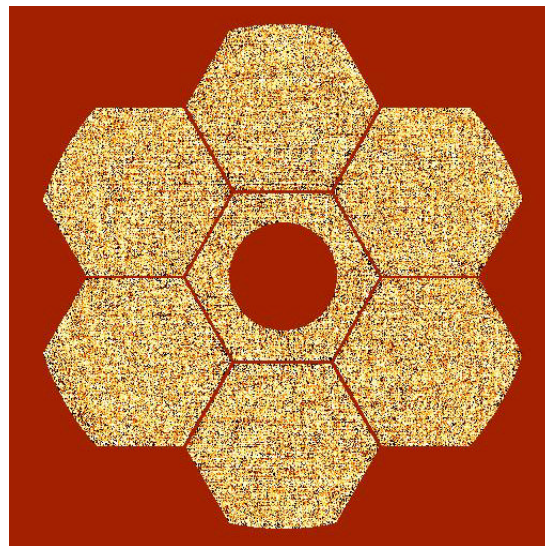
**Difference Phase
0.167 λ rms**

Figure 4.15

Phase Diversity Wavefront Results

41
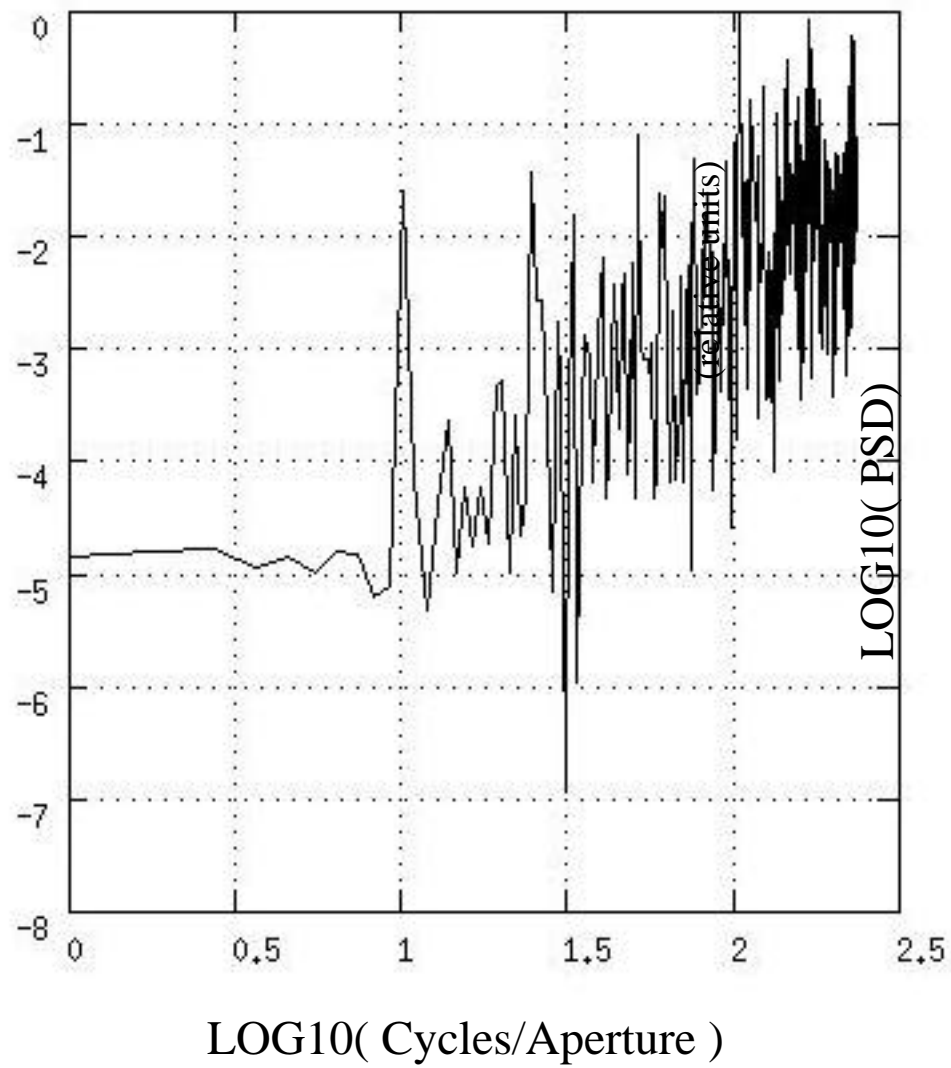
Figure 4.16 - PSD of Phase Diversity Wavefront Recovery Error
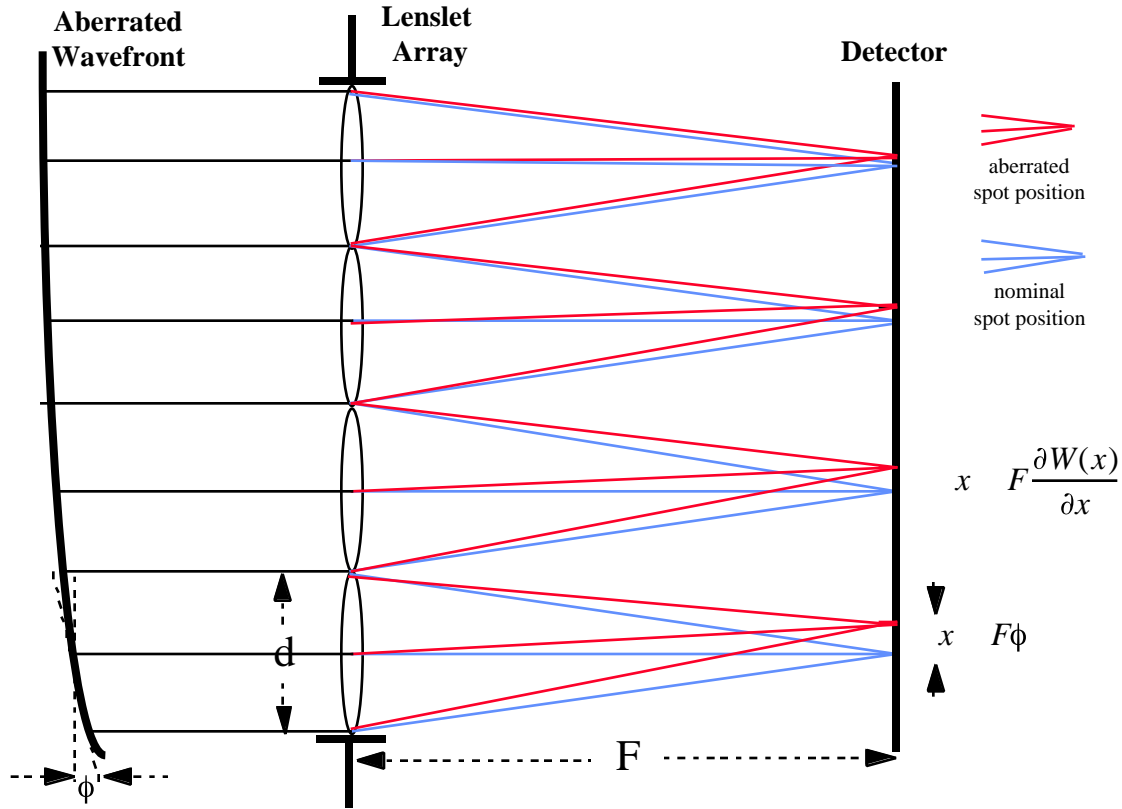
Figure 4.17
Simplified Optical Schematic of a Shack-Hartmann Wavefront Sensor

4.2 Shack-Hartmann Sensing

4.2.1 Overview, Modeling and Simulation

A diffraction based model for a Shack-Hartmann wavefront sensor has been developed and will be described in detail here. The Shack-Hartmann sensor is depicted in cartoon form in Figure 4.17. An aberrated wavefront incident on the Shack-Hartmann lenslet array is focused on the detector array. The centroid of each spot, of the grid, is shifted in the focal plane.   The shift is a linear function of the subaperture's local tilt. The lenslet array may be conveniently modeled as a phase screen which takes the incoming wavefront and adds a grid of indentations (see Figure 4.18). Each lenslet being a section of a sphere, modeled here as paraboloid. This is equivalent to assuming the spherical aberration of a single lenslet is negligible.   Thus, we model the lenslet array as:

$$E(x, y, z = 0, \lambda) = \sum_j A(x - x_j, y - y_j)e^{ikW(x,y)}e^{-i\frac{k}{2F}\left[(x-x_j)^2 + (y-y_j)^2\right]} \tag{4.2.1}$$

where "A" represents the aperture function of an individual lenslet, $W(x,y)$ represents the wavefront, (x_j, y_j) represents the position of the j-th lenslet, the sum j being taken over the entire array. Ideally one would want to propagate equation (4.2.1) to the focal plane via a Fresnel propagation evaluated by a single Fast Fourier Transform (FFT), however, this gives sampling problems. A more accurate method is to directly solve the spatial Helmholtz equation via angular spectrum methods. This is the method developed here. Equation (4.2.2), below, describes the free space propagation of the scalar electric field. The spatial Helmholtz equation is given by:

$$\nabla^2 E(x,y,z;\lambda) + k^2 E(x,y,z;\lambda) = 0 \tag{4.2.2}$$

Note that equation (4.2.2) is a 2nd order, linear, homogenous partial differential equation. Fourier transform equation (4.2.2) with respect to x and y to yield:

$$\frac{\partial^2 E(f_x, f_y; z, \lambda)}{\partial z^2} + \left[k^2 - 4\pi^2(f_x^2 + f_y^2)\right]E(f_x, f_y; z, \lambda) = 0 \tag{4.2.3}$$

This has converted equation (4.2.2) to a 2nd order ordinary differential equation with exponential solutions given by:

$$\tilde{E}(f_x, f_y; z, \lambda) = \tilde{E}(f_x, f_y; z = 0, \lambda)e^{-i\frac{2\pi}{\lambda}z\sqrt{1 - \lambda^2(f_x^2 + f_y^2)}} \tag{4.2.4}$$

This is the so called "angular spectrum" representation and is a rigorous solution to the Helmholtz equation. The term angular spectrum comes from decomposing the field into a basis of plane waves, propagating in the z-direction, at varying angles. Note that when the argument of the complex exponential becomes negative this becomes an pure real exponential representing damping at large diffraction angles i.e. evanescent waves. Thus the method of modeling the Shack-Hartmann sensor is to numerically FFT equation (4.2.1) multiply by the phasor given in equation (4.2.4) and take the

inverse FFT. The spot pattern in the focal plane of the Shack-Hartmann is then given as the modulus squared of the electric field. The leftside of Figure 4.18 shows a LEO modeled phase of 32 x 32 grid of lenslets sampled onto a 1024 x 1024 sampling grid with the DCATT pupil function, with no aberrations. This represents the null wavefront. The right side of Figure 4.18 shows a case with the topmost segment tilted and a modest amount of spherical aberration.
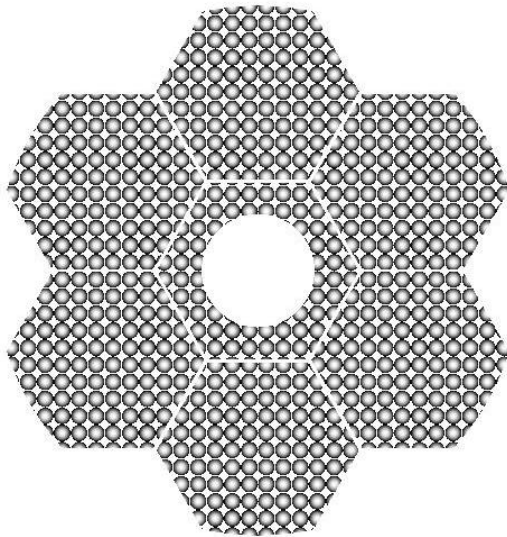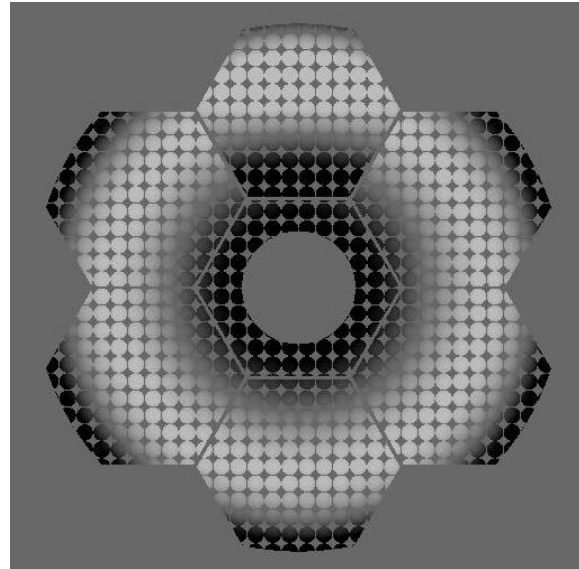
Null Wavefront   Spherical Aberration and Segment Tilt

Figure 4.18
Exit Pupil Wavefronts for 32 x 32 Shack-Hartmann Sensor

The leftside of Figure 4.19 shows the resulting focal plane spot pattern for the null wavefront case and the rightside of Figure 4.19 shows the pattern for the spherical aberration with tilt of the upper segment. Note that spots that are crossing segments and at the edge of segments are distorted. The output sampling grid is 1024 x 1024 with 9 micron focal plane sampling. The spots near the edges of the pupil also show sub-aperture spherical aberration, astigmatism, and coma due to the full aperture spherical aberration. We also simulated the phase reconstruction process. Figure 4.20 depicts this process. Shown at the top left is the spot pattern for a null wavefront (no aberrations). Shown below it the pattern with the uppermost segment tilted from left to right. Just to the right of these are shown the x and y elements of the gradients and to the right of these is the reconstructed phases. The phase reconstruction algorithm is shown as an equation between the null and the

tilted segment cases. The first thing to note is that the recovered phase is at much lower resolution than the phase retrieval process. In this case the reconstructed phase is on a 32 x 32 grid. The Shack-Hartmann approach is a natural low pass wavefront filter since very high spatial frequencies will tend to not shift the spot centroid, and thus, will not show up as aliasing. However, spatial frequencies at or near the sampling frequency of the lenslet array can be aliased. The transfer function of this low pass filter has yet to be determined. This transfer function needs to be determined to analytically compare Shack-Hartmann with Phase Retrieval. The other disadvantage of the Shack-Hartmann approach is that it cannot recover segment to segment piston errors since it measures the wavefront gradient.
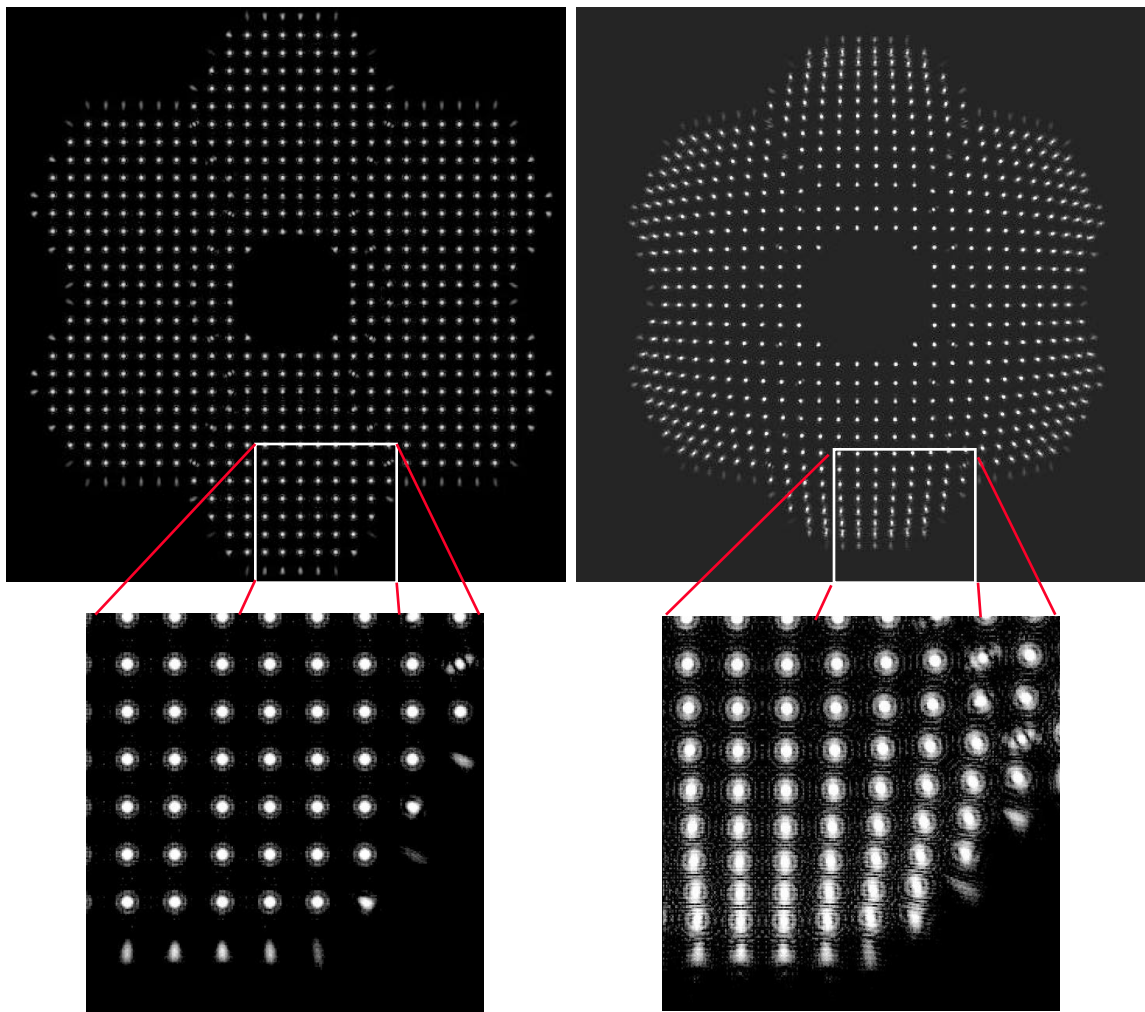
Figure 4.19
Focal Plane Spots for 32 x 32 Shack-Hartmann Sensor
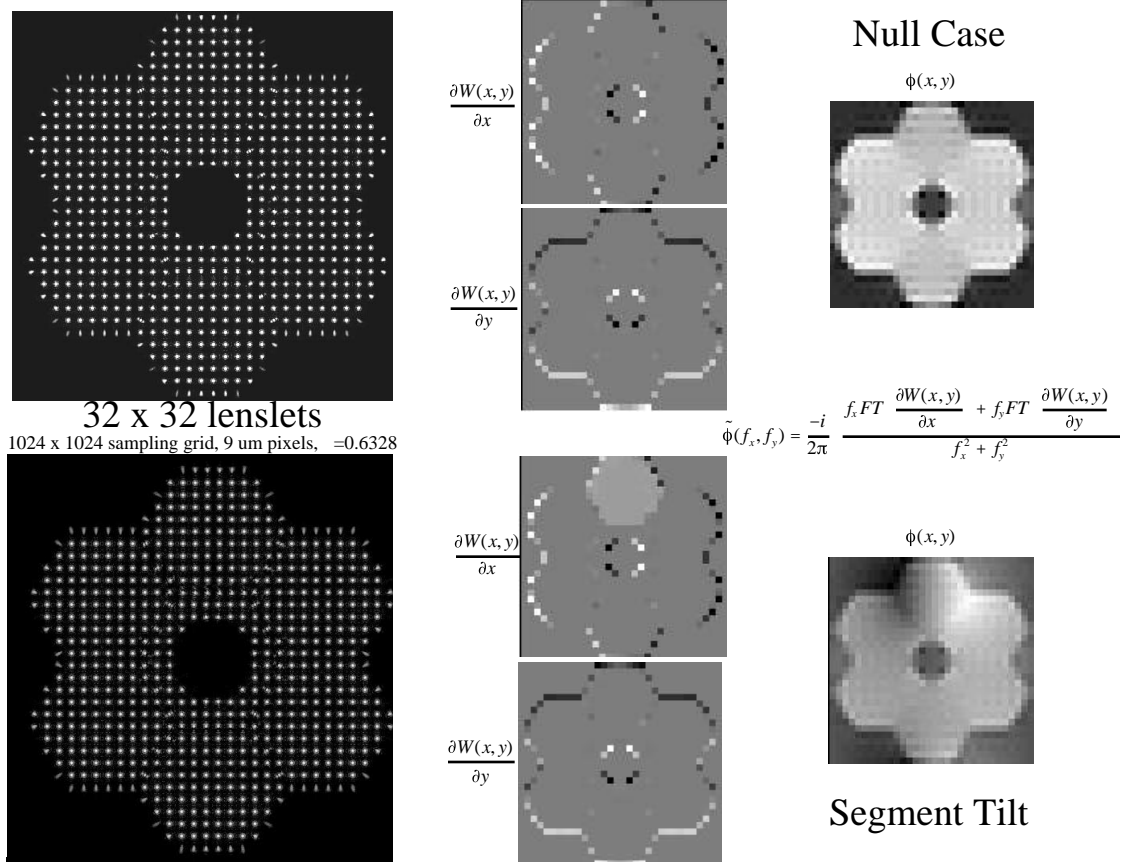
Figure 4.20

Phase Reconstruction for 32 x 32 Shack-Hartmann Array

Still to be addressed with the Shack-Hartmann sensor are the effects on WFE of:
(1) registration errors - location of the subapertures with regards to the pupil
location. (2) dynamic range - large aberrations cause spot overlaps and
confusion. The Wavescope overcomes this by grossly defocusing and then
tracking each spot until focused [34].   (3) Photon limitations, the beam is
spread out over the area of the entire pupil, thus integration times are likely
to be longer to get the same focal plane signal to noise ratio. This issues has

47

been previously addressed in previous studies [30][31], however, not specifically for DCATT.

4.2.2 Discussion and Enhancements

Ellerbroek et.al. [30] conducted a comparative study, using analysis and simulation, of Shack-Hartmann wavefront sensing and phase-diverse phase retrieval as a function of collected source photons. The lenslet samplings of the Shack-Hartmann sensor were sparse {2x2, 4x4, 8x8}. It was found that the phase diversity algorithm significantly outperforms the Shack-Hartmann for each of 3 aperture geometry's if no noise is assumed i.e., 0 read electrons. At more realistic settings of 5 read electrons rms, the phase diversity outperforms the Shack-Hartmann for photon levels above 200. However, both perform reasonably well above 10,000 photon counts. Note, these results are not directly applicable to DCATT since we would need significantly more lenslets in the Shack-Hartmann to resolve the individual segments and the sampling density for the phase diversity was low.

Rigault et.al. [31] compare Roddier's [32] curvature sensing against Shack-Hartmann sensing for the Gemini telescope. The results are from numerical simulations of the 8 meter monolithic telescope. They simulate a 9x9 and a 10x10 lenslet array case and report there results in terms of Strehl ratio. Both methods achieve nearly the same results. A Strehl of ~70% (/10) over the range of R-band magnitudes of 5 to 13 magnitudes and falling dramatically for sources with magnitudes greater than 14. R-band is centered at 0.7 um with FWHM of 0.22 um. 5 to 13 magnitudes is ~1e+8 photons to ~5e+06 photons for the aperture and detector integration times. Thus this is a relatively bright source. Note that is Strehl is obtained in simulated closed loop control with atmospheric turbulence folded in.

Roggemann and Schulz [33] develop a method to increase the dynamic range of a Shack-Hartmann sensor by using, simultaneously, a lenslet array image with a conventional image. They essentially combine a low order parametric (Zernike) phase retrieval approach to estimate the gross aberrations and combine this with the more standard Shack-Hartmann approach. The simulations shown are with primary spherical aberration and show that the dynamic range has been significantly increased. With regards to DCATT a

more interesting integration of phase retrieval and Shack-Hartmann would be to determine segment to segment piston errors. The Shack-Hartmann is blind to piston errors, phase retrieval is not. Note that the authors point out that since non-linear optimization is required the method is not well suited to high bandwidth. However, for DCATT we are essentially static, and are more interested in accuracy and precision than bandwidth. Pfund et.al, [34] develop another interesting approach to expanding the dynamic range via the use of a modified phase unwrapping algorithm combined with Shack-Hartmann wavefront sensing.

5.0 Optical Control Loop Approaches

5.1 Introduction

Telescopes such as DCATT and possibly NGST will include a deformable mirror (DM), moveable primary mirror (PM) segments and a moveable secondary mirror (SM). These moveable optical elements must be configured to give the best telescope performance for a particular set of observations, possibly including imaging, spectroscopy and coronagraphy.

Given such a system, what is the best method for determining deformable mirror, PM segment actuator movement and SM mirror rigid body movement. This is based on the choice of optical control metric and the choice of metric is based on the type of science observation required. Minimal rms wavefront error may not be the optimal choice of metric. It is generally used in ground based adaptive optical systems where high bandwidth is required. However, in DCATT and ultimately in space, we are more interested in accuracy and precision, thus, time is a lesser constraint since the atmosphere is non-existent and any system wide changes, other than jitter, will tend to be slow (~hours or ~days). In theory, jitter can be corrected with a fast steering mirror in the control loop. However this pre-supposes that jitter is well represented at time dependent wavefront tilts, i.e. rigid body motions of the PSFs which may not be the case.

Redding et.al. [1][35] have worked out the yardstick fine figure control metric for NGST which is also the baseline method for DCATT. The control method is single stage with the goal of minimizing the rms wavefront error between the phase retrieval recovered wavefront and an ideal wavefront, in this case flat. The minimal rms criterion is derived from the solution being a maximum likelihood estimator non-stationary Gaussian noise. After coarse figure control, the system hands off all DM, PM and SM actuator positions to fine figure control. The actuators may not have real positional feedback, or may have inaccuracies in positions. This vector of DM, SM and PM segment actuator information is **a**. The measured wavefront, vector **w**, is determined by a phase retrieval algorithm acting on a set of point source images taken after

coarse figure control. We currently assume a linear model of wavefront response as a function of actuator position, however, this needs to be further addressed. Thus, the wavefront-actuator model is given by:

$$\mathbf{w} = \mathbf{Ra} + \mathbf{n} \tag{5.1}$$

where $\mathbf{R}$ is the response matrix, of which the column vectors are the actuator influence functions and $\mathbf{R}_{ij} = \dfrac{\partial w_i}{\partial a_j}$; $\mathbf{n}$ represents the errors in the measured wavefront due to phase retrieval. The maximum likelihood solution, for $\mathbf{a}$, is given by:

$$\mathbf{a} = (\mathbf{R}^{t}\mathbf{C}_{v}^{-1}\mathbf{R})^{-1}\mathbf{R}^{t}\mathbf{C}_{v}^{-1}\mathbf{w} \tag{5.2}$$

Thus one moves the actuators from the current position by the amounts given by $\mathbf{a}$, i.e. $\mathbf{a}$ is the delta from the previous position. $\mathbf{C}_{v} = \langle \mathbf{nn}^{t} \rangle$ is the wavefront recovery noise covariance matrix, given as the ensemble averaged outer product of the noise vectors. It can be determined from Monte-Carlo simulations, or initially assumed diagonal unity if no information is available.

We have simulated the use of this model for a typical DM response. Our example uses 349 DM actuators within the exit pupil and a 512x512 wavefront image. Thus the response matrix, $\mathbf{R}$, has 375 columns, 349 columns of which are the DM actuator influence functions and 1 per DOF of each mirror segment and each of the 5 DOF of the secondary mirror. Thus, 375 DOF in all. Also we have the capability to turn off any arbitrary DOF and also to "clamp" or "float" a DOF. The PM actuators control tip, tilt, piston and possibly other motions for each PM segment. The columns of $\mathbf{R}$ that correspond to SM actuators must be determined uniquely for each absolute SM and PM location. These columns are calculated by ray trace methods. The ray trace calculation will yields a linear range about the current position for each SM actuator. This range is the range over which the linear model is expected to be accurate given a certain tolerance. The range will be used in the constrained optimization.

5.2 Control Methodology Flowchart

Figure 5.1 shows graphically a concept for a hierarchical fine figure control loop for DCATT. We will discuss this loop in detail here with all the options,

first we will "walk" around the loop. Starting in the upper left and proceeding clockwise. Two aberrated images, of unresolved points sources are shown, in our simulations we actually use 4 input images. These images contain the effects of random, full- and sub-aperture Zernikes, detector effects including sampling, pixelization, quantization, and noise. These images are passed into the 4 point focus diversity (Misell) phase retrieval algorithm and phase retrieved. This result, shown in the upper middle, is the phase retrieved modulo 2. The directed acyclic graph algorithm is used to unwrap the phase and the result is shown in the upper right. The resulting wavefront is least squares fit to a linear superposition of the actuator influence functions and the residual wavefront is shown in the lower middle. Note that this residual has been color stretched to enhance the low level structure, otherwise, we would see nothing but a blank frame. The resulting PSF is shown in the lower left. This shows a "walk" around the baseline control loop. In the following sections we will discuss a number of the other boxes in Figure 5.1
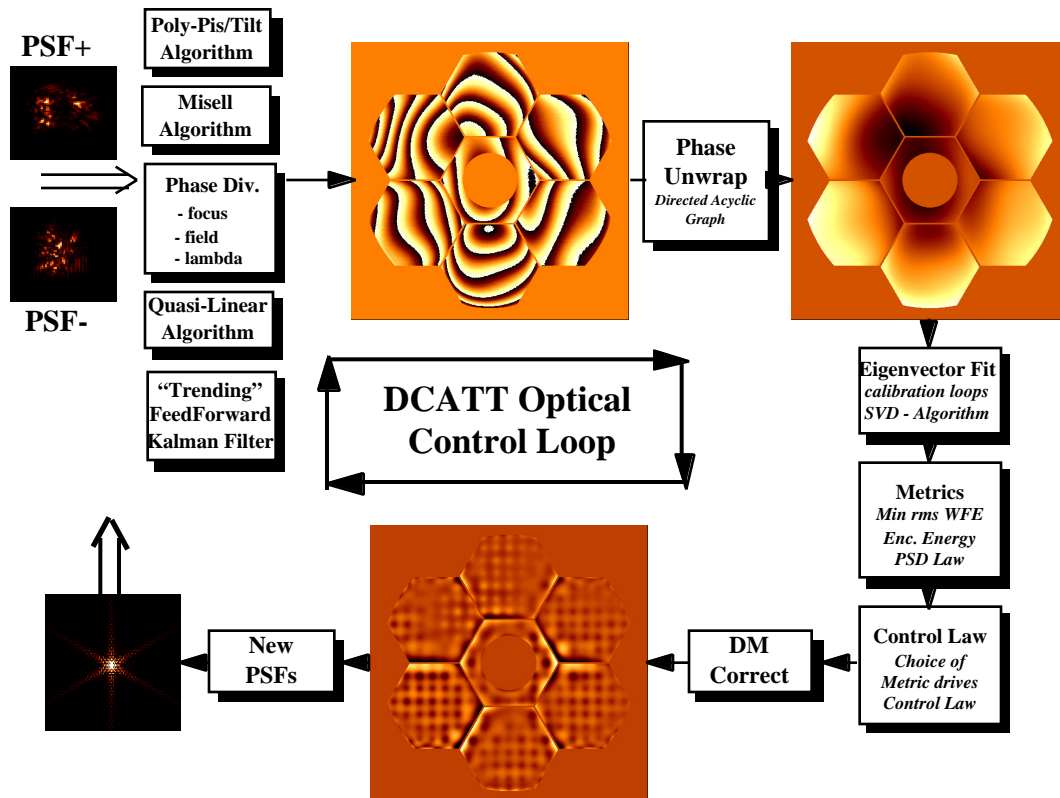

Figure 5.1
DCATT Hierarchical Optical Control Loop

5.3 Current and Future Studies

5.3.1 DM and PM Control versus Wavefront Error

In order to determine the precision with deformable mirror and primary mirror control in the active optical control loop we conducted a Monte-Carlo simulation. LEO was used to generate 100 sets of 4 PSFs, 400 PSFs in all. Each of 4 PSFs in a single set had the same realization of randomly generated full aperture Zernikes, sub aperture Zernikes including piston, tip and tilt. The rms wavefront error was varied over the range 0.0 to 2.5 waves rms WFE. The 4 PSFs in a set each had different foci of -2.0 um, -1.0 um, +1.0 um, +2.0 um rms WFE. The PSFs were polychromatic with a center wavelength of =0.6328 um with a FWHM of =10 nanometers. The detector MTF as well as 512 x 512 sampling grids were folded in. Each of the 100 sets were passed through the 4 point focus diversity algorithm and the phase retrieved. The results were phase unwrapped and fit to actuator influence functions for the DM and fit to tip/tilt/piston terms for the PM. The results for the WFS precision was previously plotted in Figure 4.6 The reconstructed wavefront was subtracted from the phase retrieved wavefront to simulated the correction. This residual wavefronts standard deviation was plotted versus input RMS wavefront error in Figure 5.2 for 3 separate cases. The "redline" in Figure 5.2 represents the RMS residual in units of waves of wavefront error for both DM and PM control, but, with no range limits. The "blueline" represents both DM and PM control but with DM range limits and the "greenline" represents DM control only. Note that the "blueline" is the more correct case since the DM does have limitations on it's range. The control with range limited DM and PM is approximately /10 for an input WFS error of upto 1.0 waves RMS. Note that this simulation assumes that the system is static between successive iterations of the control loop. This implies that nothing would be gained by continued iteration. No more wavefront error could be corrected. The residual wavefront error comes from structure and spatial frequencies which cannot be corrected with this configuration of actuators.
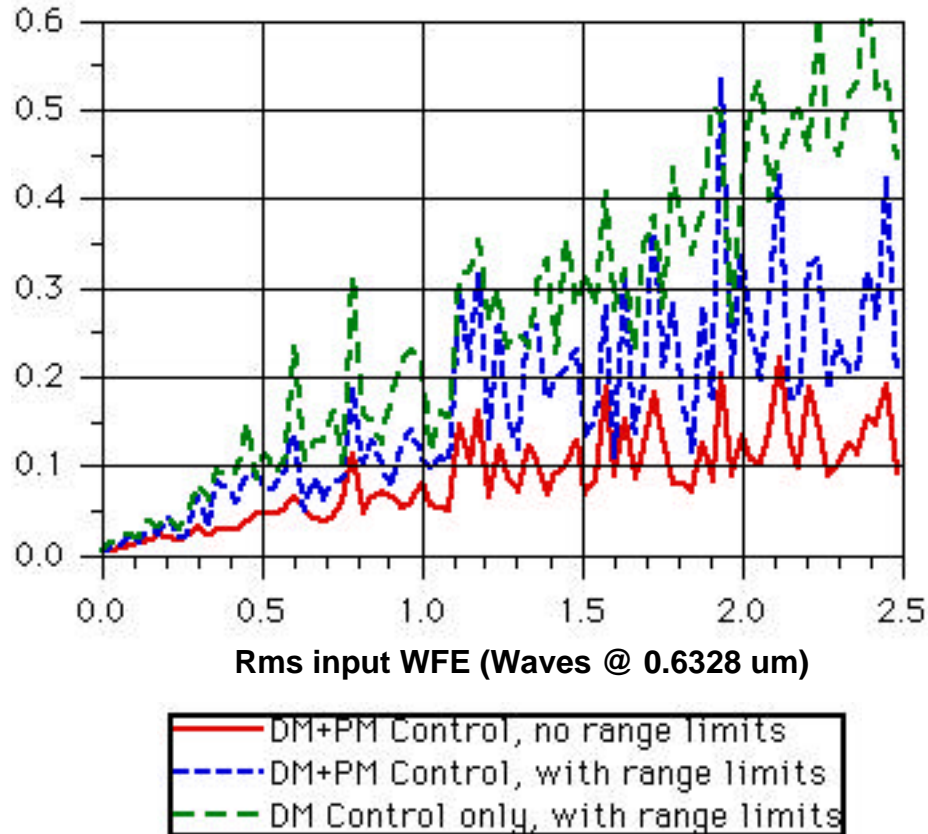
Figure 5.2
DCATT Control Loop Simulations

5.3.2 Error Budget Modeling

The simulations delineated in this report have been summarized in Table 5.1.

Over the budgeted range of 0.0 - 1.0 waves of RMS wavefront error we expect wavefront sensing to contribute ~ /23, DM+PM control ~ /10, jitter ~ /8, SNR ~ /10, root sum squaring these together yields an RSS of /5.2 which compares favorably with the budgeted item of /4.43. It appears from the modeling and simulation that the error budget can be met.

Over budgeted range 0 - 1 λ rms input WFE for DCATT

WFS **DM+PM Control** **Jitter** **SNR**	$\sim\!\lambda$**/23** **< λ/10** **< λ/8 (0.25 rms jitter)** **~ λ/10 (Fullwell > 20000, SNR > 62**
σ **(RSS)** σ **(error budget)**	$\sim\!\lambda$**/5.2** λ**/4.43**

Table 5.1
DCATT Control Loop Simulation Error budget

5.3.3 Control Metrics and Image Quality

Strehl ratio is the ratio of the peak intensity at the focal point of the actual aberrated system to the peak intensity at the same point for a perfect, unaberrated system.   The maximal Strehl ratio occurs for a minimal rms wavefront error. The squared wavefront error, σ_w^2, is defined by:

$$\sigma_w^2 = (\mathbf{w} - \mathbf{Ra})^T (\mathbf{w} - \mathbf{Ra}) \tag{5.3}$$

where $\mathbf{Ra}$ is the reconstructed wavefront and $\mathbf{w}$ the true wavefront error, not that which is returned from phase retrieval. Given an initial error wavefront, $\mathbf{w}$, the minimum unconstrained wavefront error is given by:

$$\mathbf{a} = \left(\mathbf{R^t R}\right)^{-1} \mathbf{Rw} \tag{5.4}$$

This is the same condition as maximizing the Strehl ratio. Optimizing for the least-squared error has the advantage of using a straight forward calculation, that is rapidly computed. One disadvantage is that use of the DM can leave the wavefront with a dimpled figure, having low mean-squared error but much of the error in higher spatial frequency components. This is simply because of the sampling constraints required by the actuator density.

An alternative metric would be to maximize encircled energy within a radius of the PSF. This is generally more optimal for spectroscopic observations. It is believed this may reduce some of the sidelobes in the PSF at the expense of a reduced Strehl ratio. In this study we have not yet attempted to utilize this as a metric but this is a recommendation to worth further research and subsequent testing on DCATT. As an alternative metric for coronagraphic observations, it may be best to minimize, or roll-off, high spatial frequency variations in the wavefront, thereby reducing the probability of a PSF side-lobe appearing as a false star. This would also tend to reduce the Strehl ratio and give a slightly broader PSF core. The problem becomes.

$$Minimize\ S(f_x, f_{y)})PSD_r(f_x, f_y ; \mathbf{a}) \qquad (5.5)$$

Where **S** is the penalty weighting function and PSD is the Power Spectral Density of the corrected wavefront as a function of the actuator vector **a**. Both the encircled energy and the coronagraphic metrics would become nonlinear optimization problems as opposed to the linear least squares metric, i.e., minimal RMS wavefront error. The disadvantage of both of these metrics are that they are more computationally complex, thus, taking more time. This may not be a problem for DCATT, since it is essentially a static system. It may not be a problem for NGST.

5.3.4 Actuator Constraints

For the least-squared and least-summed error metrics, equation (5.3) is optimum for the unconstrained problem. If, however, the solution, **a**, violates one or more physical constraints of the system, a constrained optimization solution must be found.

The constraints include:
1. Upper and lower travel limits on all actuators.
2. DM slope constraints, No nearest neighbor DM actuators can differ by more than 1 micron.
3. One or more actuator can fail, possibly being frozen in place or be free to float, also there are usable actuators outside the pupil boundaries, which may need to be weighted in the fit.

This type of constrained optimization problems is referred to as linear programming, when the cost function to be minimized is a linear function of the unknown, and quadratic programming, when the cost function is quadratic. Currently in our simulations if an actuator is range limited then it's value is just truncated to the range limit. This is sub-optimal, however, we recommend looking at the constrained optimization problem in more detail since a better solution most certainly will be found for the minimal RMS wavefront error metric. This research would also be required for both the encircled energy and power spectral density metrics.

5.3.5 System Identification

Any drift in actuator influence functions lowers achievable image quality. For NGST the changes may be due to gravity release, material property change, thermal drift and even human error. Thus it behooves us to look at ways of in-situ re-calibration of the actuator influence functions on-orbit. Thus, we recommend that DCATT try a series of experiments to determine how well this can be done with the DCATT system. A series of error could be introduced into the influence function matrix and the system identification algorithm such as the following could be tried. We would also need to research the different methods of system identification further and could also model and simulate this. One system identification algorithm [36] is shown below. For the wavefront model given by:

$$\mathbf{w} = \mathbf{R}\mathbf{a} + \mathbf{e} \tag{5.6}$$

where $\mathbf{e}$ represents a vector of errors. The influence function matrix can be updated according to the following algorithm:

$$\mathbf{R}_{(n+1)} = \mathbf{R}_{(n)} + \left[\mathbf{w}^{(n+1)} - \mathbf{R}_{(n)}\mathbf{a}^{(n+1)}\right]\mathbf{K}_{(n+1)} \tag{5.7}$$

$$\mathbf{K}_{(n+1)} = \frac{\mathbf{a}^{t(n+1)}\mathbf{P}_{(n)}}{1 + \mathbf{a}^{t(n+1)}\mathbf{P}_{(n)}\mathbf{a}^{(n+1)}} \tag{5.8}$$

$$\mathbf{P}_{(n+1)} = \mathbf{P}_{(n)} - \frac{\mathbf{P}_{(n)}\mathbf{a}^{(n+1)}\mathbf{a}^{t(n+1)}\mathbf{P}_{(n)}}{1 + \mathbf{a}^{t(n+1)}\mathbf{P}_{(n)}\mathbf{a}^{(n+1)}} \tag{5.9}$$

$$\mathbf{P}_{(n)} = \left(\mathbf{A}_{(n)}\mathbf{A}^{t}_{(n)}\right)^{-1} \tag{5.10}$$

and $\mathbf{A}_{(n)}$ is a matrix containing all the previous fits to the actuator influence functions. Note that this method is computationally complex and still needs to be researched further.

5.4 Summary

We have begun looking at some control aspects of DCATT, and NGST, but strongly recommend that much more effort be put into this. Redding et.al. [1] have made a great beginning to this important aspect of DCATT, however, there are still many open issues. The control aspects is one area where high fidelity modeling and simulation can play a major role. Some of the open issues that modeling and simulation could play a role are:

(1) Optimal approach to calibration of the control loop ? The influence functions will be measured on the optical bench before insertion into DCATT, but, is there a good way to measure the influence functions in the system ? A simple method, but very time consuming, would be just to move each actuator in turn, a unit amount, observe the change in wavefront through the wavefront sensing. However this is very time consuming requiring one, or more, cycle(s) of the entire control loop for each actuator. Alternatively we could perform multiple actuators movements since actuators which are not local to each other will exert little influence on each other. A more optimal way would be to perform an eigenvalue decomposition of the system response matrix and sort on descending eigenvalues. This would give orthogonal modes, i.e., linear combinations, of actuators which would not influence each other. Sorting on the eigenvalues gives those modes with the most gain. The highest gain modes should be calibrated first.

(2) In-situ re-calibration, we need to further investigate both in-situ calibration for DCATT, model a number of the methods for the different possible control metrics and conduct a series of experiments on DCATT to validate the predicted performance of these methods. I consider this a relatively critical step since any error in our knowledge of NGST's influence functions will adversely affect wavefront and thus image quality.

(3) Trending, feed forward and Kalman filtering. We could deliberately introduce known, slowly varying thermal gradients into the DCATT primary

mirror and obtain wavefront trend data. This data could be used to develop a predictive feed forward controls approach which could be periodically updated with observed data.

(4) Mode visibility problem, The optical system design and actuators can introduce aberrations into the system which cannot be adequately sensed by the wavefront sensing method, and, thus cannot be corrected by closed loop control. Reference [37] discuss this problem in more detail and how to determine if this problem exists and how to circumvent it.

6.0 Recommendations

DCATT Data

* Standardize all data file formats, e.g., FITS format, including all calibration and datafiles.

* Insure tracebility of all calibration files and data files, e.g. use time tags and descriptive headers.

* Save and archive all data runs, even failed runs.

* Make all data available to any and all teams via a web-based interface. This includes all government, industry and university.

* Furthermore, a guest investigator program should be instituted to allow outside investigators to test and demonstrate their methods.

DCATT Modeling Efforts

* Continue development of independent models for DCATT.

* Cross validate these independent models.

* Validate independent models with real data.

* Identify model deficiencies and differences.

* If possible correct deficiencies and understand differences.

NGST Scalability and Model Extrapolation

* Identify relevant scalability issues to NGST.

* Which issues can and cannot be addressed by DCATT ?

* What modifications could be done to address those not currently addressed ?

* Will DCATT validated models be able to predict NGST performance ? will we be able to believe the predictions ?

* Simulate spectrum and system response of NGST jitter. How do the wavefront sensing models perform with this jitter ? Can we introduce jitter of this type into DCATT ?

Wavefront Sensing

- Further investigate phase diversity as an in-situ wavefront quality monitor. Investigate field diversity, wavelength diversity and object diversity.

- Are other fast image based metrics for wavefront quality on extended sources available ?

- Study and potential coupling effects of the fast steering mirror into phase retrieval.

- Look further into inteferometric methods and Shack-Hartmann methods. Although Shack-Hartmann cannot determine piston it is very fast and simple in comparison with phase retrieval based methods.

- Investigate quasi-linear phase retrieval methods [38]; these are computationally simpler but have limited dynamic range and other constraints.

Optical Controls

- Simulate and attempt system identification methods on DCATT to identify the system response to compensate for drift.

- Determine the optimal wavefront quality subject to the constraints imposed by the actuators and deformable mirror. This takes the mathematical form of a constrained optimization.

- Investigate the effects of other control metrics such as maximum encircled energy and/or power spectral density slope. The current metric is based on minimal rms wavefront error. This may not be the best metric for astronomical observations.

- What is the optimal method to initially calibrate the system.

- Determine the need and architecture for hierarchical control strategies.

7.0 Summary and Conclusions

We have presented the results of an on-going study for comparative wavefront sensing for the Developmental Comparative Active Telescope Testbed. In this report we have discussed our independent DCATT systems model, including segmented apertures, raytracing, aberrations, misalignments, residual polish marks, scattering, jitter, detector effects and noise. We have begun a comprehensive comparison of a focus diverse Misell phase retrieval algorithm, a wavelength diverse phase retrieval algorithm, an extended object phase retrieval algorithm and a Shack-Hartmann slope measuring wavefront sensor. We have also discussed a hierarchical control loop, and calibration of the control. This is an area where much further study is recommended. This study is entirely based upon analytical/mathematical models, high fidelity computer models, sensitivity analysis and Monte-Carlo type studies. At the time of this writing no actual data is available through the DCATT telescope and active optical testbed, however, it will shortly become available.

8.0 References

[1] Redding, D., Basinger, S., Lowman, A., Kissil, A., Bely, P., Burg, R.,Mosier, G., Femiano, M., Wilson, M., Schunk, G., Craig, L., Jacobson, D., Rakoczy, J., Hadaway, J., "*Wavefront Sensing and Control for a Next Generation Space Telescope*", Proceedings of SPIE Vol. 3356, Kona Hawaii, March 1998

[2] Lyon, R. G., Hollis, J. M, Dorband, J., Murphy, T.P., "*Extrapolating HST Lessons to NGST*", Optics and Photonics News, July 1998, Vol. 9, No 7

[3] Seery, B.D.,*"Next Generation Space Telescope (NGST)"*, Proceedings of SPIE, Vol. 3356, March 1998, Kona Hawaii

[4] Mather, J.C., Stockman, P., *"NGST: Technology for Looking Back in Time"* , BAAS., **191**, #54.02, Dec 1997

[5] Seery, D. and Smith, E. P., "*NASA's Next Generation Space Telescope: Visiting a Time When Galaxies Were Young*", Optics & Photonics News, July 1998, pp. 29-34.

[6] Lebeouf, C. M., Davila, P.S., Redding, D.C., Morell, A., Lowman, A.E., Wilson, M.E., Young, E.W., Pacini, L. K.,Coulter D.R.,*"Deployed Cryogenic Active Telescope Testbed for the Next Generation Space Telescope",* Proceedings of SPIE, Space Telescopes and Instruments V, 3356, March 1998, Kona Hawaii

[7] Pacini, L., Geithner, P.H., Jacobson, D.N., Redding, D.C.,*" NEXUS: The Next Generation Space Telescope Integrated Flight Systems Demonstration*", Proceedings of SPIE, Vol. 3356, March 1998, Kona Hawaii

[8] Davila, P.S., Lowman, A.E., Wilson, M. E., Boucarut R. A., LeBoeuf, Redding D.C., Young, E.W., *"Optical Design of the Developmental Cryogenic Active Telescope Testbed",* Proceedings of SPIE, Space Telescopes and Instruments V, **3356**, 1998, Kona HI.

[9] Young, E.W., Wilson, M.E., Davila, P.S., Eichhorn W., LeBoeuf, C.M., Redding, D.C., Lowman, A.E., *"Performance Analysis of the Developmental Cryogenic Active Telescope Testbed (DCATT)"*, Proceedings of SPIE, Space Telescopes and Instruments V, **3356**, 1998, Kona HI.

[10] L. Grey, R.G. Lyon, "*Correction of Misalignment Dependent Aberrations of the Hubble Space Telescope via Phase Retrieval*", Proceedings of SPIE, Vol. 1168, August 1989

[11] Lyon, P.E. Miller, "Phase Retrieval Algorithms and Results", in *Proceedings of the First Hubble Aberration Recovery Program Workshop*, November 15-16, 1990, R. Korechoff ed. (Jet Propulsion Laboratory, Pasadena, CA) paper #13

[12] Lyon, P.E. Miller, A. Gruszczak, "*HST Phase Retrieval Special Studies Task - Final Report*", HDOS Report PR J14-0013, NASA contract NAS 8-38494, Feb 1991

[13] Lyon, P.E. Miller, A. Gruszczak, "*Hubble Space Telescope* Phase Retrieval: A Parameter Estimation", Proceedings of SPIE, Vol. 1567, July 1991

[14] Lyon, R. G., Dorband, J. E., Hollis, J. M., "*Hubble Space Telescope Faint Object Camera Calculated Point Spread Functions*", Applied Optics, Vol. 36, No. 8, March 10, 1997.

[15] W. Goodman, " *Introduction to Fourier Optics* ", McGraw-Hill Inc., 1986

[16] R.W. Gerchberg, W.O. Saxton, *"A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures",* Optik, **35**, 237-246

[17] Misell, *"A Method for the Solution of the Phase Problem in Electron Microscopy"*, J. Phys. D, Vol 6, L6-L9, 1973

[18] Baba, H. Tomita, N. Miura, "*Iterative Reconstruction Method in Phase Diverse Imaging*", Applied Optics, **33**, No. 20, (1994)

[19] J. Fienup, *" Phase Retrieval for the Hubble Space Telescope Using Iterative Propagation Algorithms",* Proceedings of SPIE, Vol. 1567, July 1991

[20] R.A. Gonsalves, "*Phase Retrieval and Diversity in Adaptive Optics*", Opt. Eng. **21**, 829 (1982)

[21] R.G. Paxman, and S. L. Crippen, "*Aberration Correction for Phased-Array Telescopes using Phase Diversity*", in Digital Image Synthesis and Inverse Optics, Proceedings of SPIE **1351**, pg. 787-797, (1990)

[22] R.G. Paxman, and J.R. Fienup, "*Optical Misalignment Sensing and Image Reconstruction using Phase Diversity*", JOSA, **5**, 914-923 (1988)

[23] R.L. Kendrick, D.S. Acton, and A.L. Duncan, "*Phase-Diversity Wavefront Sensor for Imaging Systems*", Applied Optics, Vol. 33, No. 27, September 20,1994

[24] R.L. Kendrick, D.S. Acton, and A.L. Duncan,"*Experimental Results from the Lockheed Phase Diversity Test Facility*", Proceedings of SPIE, Vol 2302, Image Reconstruction and Restoration, July 25-26, 1994

[25] Lee, M.C. Roggemann, B.M. Welsh, E.R. Crosby, "*Evaluation of Least-Squares Phase Diversity Technique for Space Telescope Wavefront Sensing*", Applied Optics, **36**, No. 35, December 10, 1997

[26] R.G. Paxman, J.H. Seldin, M.G. Lofdahl, G.B. Scharmer and C.U. Keller, "*Evaluation of Phase-Diversity Techniques for Solar-Image Restoration*", Astrophysical Journal, **466**, 1087-1099, August 1, 1996

[27] M. G. Lofdahl,"*Phase Diversity Wavefront Sensing and Image Reconstruction Applied to High-Resolution Solar Observations*", Doctoral Dissertation, University of Stockholm, S-10691, 1996

[28]J.H. Seldin and R. G. Paxman, " *Phase-Diverse Speckle Reconstruction of Solar Data*", Proceedings of SPIE, **2302-19**, San Diego, CA, (1994)

[29] R.G. Paxman, T.J. Schulz and J.R. Fienup, "*Joint Estimation of Object and Aberrations by using Phase Diversity*", JOSA, **9**, No. 7, pp. 1072-1085, (1992)

[30] B.L. Ellerbroek, B.J. Thelen, D.J. Lee, D.A. Carrara and R.G. Paxman, "*Comparison of Shack-Hartmann Wavefront Sensing and Phase Diverse Phase Retrieval*", Proceedings of SPIE, Adaptive Optics and Applications, **3126**, San Diego CA, July 1997

[31] F. Rigaut, B.L. Ellerbroek, M.J. Northcott, "*Comparison of Curvature-Based and Shack-Hartmann-based Adaptive Optics for the Gemini Telescope*", Applied Optics, **36**, No. 13, May 1, 1997

[32] F. Roddier, C. Roddier, and N. Roddier, "*Curvature Sensing: a new Wavefront Sensing Method*", in Statistical Optics, G.M. Morris, ed. Proceedings of SPIE, **976**, 203-209 (1988)

[33] M. C. Roggemann, T.J. Schulz, "*An Algorithm to increase the largest aberration which can be reconstructed from Hartmann Sensor Measurements*", Applied Optics,**37**, No. 20, July 10, 1998

[34] A. Wirth, A. Jankevics, F. Landers, C. Baird and T. Berkopac, "Final Report on the testing of the CIRS telescopes using the Hartmann Technique", Technical Report NAS5-31786, Adaptive Optics Associates, Cambridge, MA, (1993)

[35] Lau, Kenneth, Breckenridge, Bill, Norheim, Noble, and Redding, David, "*Active Figure Maintenance Control using an Optical Truss Laser Metrology System for a Space-Based Far-IR Segmented Telescope*", SPIE , Controls of Optical Systems, Vol. 1696, 1992, pp. 61-82.

[36] B.D.O. Anderson, "*An Approach to Multivariable System Identification*", Automatica, **13**, No. 13, pp. 401-408, (1977)

[37] M.E. Furber and D. Jordan, "*Optimal Design of Wavefront Sensors for Adaptive Optical Systems: Part 1, Controllability and Observability Analysis*", Optical Engineering, **36**, 7, pp 1843-1855 (1997)

[38] W. J. Wild, "Linear Phase Retrieval for Wavefront Sensing", Optics Letters, **23**, issue 8, April 15, 1998

Appendix A
Table A.1: *Sample LEO Aperture Input File*

```
POLY = 1 {
        Nsides =              6          # Number of sides of polygon
        radius =  0.173205               # radius of inscribed polygon (meters)
        xcent =   0.000000               # X-center of polygon (meters)
        ycent =   0.000000               # Y-center of polygon (meters)
        theta =   0.000000               # rotation angle of polygon (degrees))
        piston =  0.000000               # piston error (microns-surface error)
        xtilt =      0.000000            # tilt of mirror in x-direction (arcsec)
        ytilt =      0.000000            # tilt of mirror in y-direction (arcsec)
        Z01   =  0.066854                # Piston (microns WFE )
        Z02   =  0.057037                # X-tilt (microns WFE )
        Z03   =  0.016728                # Y-tilt (microns WFE )
        Z04   = -0.025847                # X-Y astigmatism (microns WFE )
        Z05   =  0.044261                # Focus (microns WFE )
        Z06   =  0.001666                # 45-degree astigmatism (microns WFE )
        Z07   = -0.013957                # Trefoil (microns WFE )
        Z08   = -0.050279                # X-coma  (microns WFE )
        Z09   =  0.064685                # Y-coma  (microns WFE )
        Z10   = -0.034657                # Trefoil (microns WFE )
        Z11   =  0.000232                # (microns WFE )
        Z12   =  0.065174                # (microns WFE )
        Z13   = -0.048524                # Fourth-order spherical (microns WFE )
        apodize =            N           # anti-alias mask (Y/N).
        }
POLY = 2 {
        Nsides =              6          # Number of sides of polygon
        radius =  0.173205               # radius of inscribed polygon (meters)
        xcent =   0.000000               # X-center of polygon (meters)
        ycent =   -0.306000              # Y-center of polygon (meters)
        theta =   0.000000               # rotation angle of polygon (degrees))
        piston =  0.000000               # piston error (microns-surface error)
        xtilt =      0.000000            # tilt of mirror in x-direction (arcsec)
        ytilt =      0.000000            # tilt of mirror in y-direction (arcsec)
        Z01   =  0.042932                # Piston (microns WFE )
        Z02   =  0.067594                # X-tilt (microns WFE )
        Z03   = -0.050003                # Y-tilt (microns WFE )
        Z04   = -0.022361                # X-Y astigmatism (microns WFE )
        Z05   =  0.027009                # Focus (microns WFE )
        Z06   = -0.004060                # 45-degree astigmatism (microns WFE )
        Z07   = -0.011545                # Trefoil (microns WFE )
        Z08   = -0.025496                # X-coma  (microns WFE )
        Z09   =  0.023511                # Y-coma  (microns WFE )
        Z10   = -0.012937                # Trefoil (microns WFE )
        Z11   =  0.001285                # (microns WFE )
        Z12   = -0.031798                # (microns WFE )
        Z13   =  0.064121                # Fourth-order spherical (microns WFE )
        apodize = N                      # anti-alias mask (Y/N).
        }
•
•     through segment number 7
•
POLY = 7 {
        Nsides =              6          # Number of sides of polygon
        radius =  0.173205               # radius of inscribed polygon (meters)
        xcent =  -0.265004               # X-center of polygon (meters)
        ycent =  -0.152000               # Y-center of polygon (meters)
        theta =   0.000000               # rotation angle of polygon (degrees))
        piston =  0.000000               # piston error (microns-surface error)
        xtilt =      0.000000            # tilt of mirror in x-direction (arcsec)
        ytilt =      0.000000            # tilt of mirror in y-direction (arcsec)
        Z01   =  0.060268                # Piston (microns WFE )
        Z02   = -0.016569                # X-tilt (microns WFE )
        Z03   =  0.027167                # Y-tilt (microns WFE )
        Z04   = -0.004950                # X-Y astigmatism (microns WFE )
        Z05   = -0.059248                # Focus (microns WFE )
        Z06   = -0.032692                # 45-degree astigmatism (microns WFE )
        Z07   =  0.000516                # Trefoil (microns WFE )
        Z08   = -0.038527                # X-coma  (microns WFE )
        Z09   = -0.045402                # Y-coma  (microns WFE )
        Z10   = -0.041302                # Trefoil (microns WFE )
        Z11   =  0.006547                # (microns WFE )
        Z12   =  0.031134                # (microns WFE )
        Z13   = -0.057388                # Fourth-order spherical (microns WFE )
        apodize = N                      # anti-alias mask (Y/N).
        }
```

Table A.2
Sample LEO Diffraction Input File

```
=          0.903              # Exit pupil Diameter (meters)  (system is F/15)
=          0.6328             # Wavelength (microns)
=          0.1330311          # Output Sample spacing (arseconds) (pixels are 9 um)
=          Y                  # Generate PRF (Y or N)
=          0.1330311          # Detector element size for PRF (arseconds)
Npsf= 7                       # Number of PSF's across Passband.
fwhm= 0.0100                  # FWHM of Filter (microns) (if Npsf's > 1 )
=          0.21384            # aperture obscuration ratio (0.21384)
Ztype= 1                      # Zernike type (0 = Annular, 1 = Code 5)
 Z01= 0.083568     # (microns WFE )
 Z02= 0.071296     # (microns WFE )
 Z03= 0.020909     # (microns WFE )
 Z04= -0.032309    # (microns WFE )
 Z05= -0.944674    # (microns WFE )
 Z06= 0.002082     # (microns WFE )
 Z07= -0.017446    # (microns WFE )
 Z08= -0.062849    # (microns WFE )
 Z09= 0.080856     # (microns WFE )
 Z10= -0.043321    # (microns WFE )
 Z11= 0.000290     # (microns WFE )
 Z12= 0.081467     # (microns WFE )
 Z13= -0.060655    # (microns WFE )
 Z14= -0.051053    # (microns WFE )
 Z15= -0.025705    # (microns WFE )
 Z16= 0.067937     # (microns WFE )
 Z17= -0.001606    # (microns WFE )
 Z18= 0.034743     # (microns WFE )
 Z19= 0.008727     # (microns WFE )
 Z20= 0.077236     # (microns WFE )
 Z21= -0.017989    # (microns WFE )
 Z22= -0.056010    # (microns WFE )
 Z23= 0.028798     # (microns WFE )
 Z24= -0.054851    # (microns WFE )
 Z25= 0.007807     # (microns WFE )
 Z26= 0.063089     # (microns WFE )
 Z27= 0.059521     # (microns WFE )
 Z28= 0.069274     # (microns WFE )
 Z29= 0.057818     # (microns WFE )
 Z30= -0.046393    # (microns WFE )
 Z31= 0.026656     # (microns WFE )
 Z32= 0.052583     # (microns WFE )
=    0.000000                 # OTA aperture rotation angle (degrees)
=    0.0                      # x - Jitter (milli-arcseconds)
=    0.0                      # y - Jitter (milli-arcseconds)
=    Y                        # Create  phase map file ?(Y or N)
=    N                        # Use Recovered Surface Map (Y/N/P),if Y or P then PM=SM=N.
=    N                        # Apodized the Pupil function ? (Y or N)
=    N                        # Add in random Gaussian surface (Y or N)
=    0.01                     # S. Dev. of random Gaussian surface (microns).
=    993                      # integer seed value for random surface generator.
=    act0.in                  # Actuator file ("none" if not using)(units are um WFE).
Nact= 349                     # Total number actuators in pupil.
act_cof= 39.8175              # R(r) = exp(-act_cof*r)*sin(act_cof*r+PI/4)
fullwell  = 80000.0           # Detector fullwell in electrons.
shot     = Y                  # Add in shot noise (Y/N).
readnoise = 13.0              # Detector readnoise in electrons.
quant    = 4096.0             # number of quantize levels (4096 levels, 0=>not quantize).
```

Appendix B
Mathematical Derivation of Diamond Turned Scattered Light on PSF

Approach 1: Linearized Scattering Theory

The linearized scattering approach is based on both the small angle approximation and prior knowledge of the statistical properties of the surface. The ensemble averaged monochromatic diffraction point spread function for an imaging system is given by:

$$PSF(u,v) = \left\langle \left| \frac{1}{\lambda^2 F^2} \int_{aper} A(x,y) e^{i\frac{2\pi}{\lambda}W(x,y)} e^{-i2\pi(\frac{xu}{\lambda F}+\frac{yv}{\lambda F})} \, dxdy \right|^2 \right\rangle \qquad (B.1)$$

where x and y are the exit pupil plane coordinates, u and v the focal plane coordinates. A(x,y) is a zero or one function, representing the support of the pupil function, i.e. zero where no light passes and unity where it does. F is the system focal length and the wavelength. W(x,y) is the exit pupil wavefront.

If we split the wavefront, W(x,y), into 2 components given by:

$$W(x,y) = W_d(x,y) + W_s(x,y) \qquad (B.2)$$

where $W_d(x,y)$ represents the deterministic wavefront, e.g. due to focus, and $W_s(x,y)$ represents the non-determinstic wavefront given by $W_s(x,y) = 2S(x,y)$ where S(x,y) is mirror surface function responsible for scattered light from the mirror surfaces. Furthermore, assume that the scattered wavefront is a zero mean random process such that:

$$\langle S(x,y) \rangle = 0, \quad \text{and} \quad \langle S^2(x,y) \rangle = \sigma_s^2 \quad \text{and} \quad S(x,y) << \lambda \qquad (B.3)$$

Thus we can expand the exit pupil phasor in equation (B.1) to be:

$$e^{i\frac{2\pi}{\lambda}W(x,y)} \qquad e^{i\frac{2\pi}{\lambda}W_d(x,y)} \qquad 1 + i\frac{2\pi}{\lambda}W_s(x,y) \qquad (B.4)$$

and equation (B.1) becomes:

$$PSF(u,v) \quad \left\langle \frac{1}{\lambda^2 F^2}\bigg|_{aper} A(x,y)e^{i\frac{2\pi}{\lambda}W_d(x,y)} \left[1 + i\frac{2\pi}{\lambda}2 S(x,y)\right] e^{-i2\pi(\frac{xu}{\lambda F}+\frac{yv}{\lambda F})} dxdy\bigg|^2 \right\rangle \quad (B.5)$$

Thus we can seperate equation (B.5) into 2 integrals of the form:

$$PSF(u,v) \quad \frac{1}{\lambda^2 F^2}\left\langle \bigg| \int_{aper} A(x,y)e^{i\frac{2\pi}{\lambda}W_d(x,y)} e^{-i2\pi(\frac{xu}{\lambda F}+\frac{yv}{\lambda F})}dxdy + i\frac{4\pi}{\lambda}\int_{aper} A(x,y)e^{i\frac{2\pi}{\lambda}W_d(x,y)}S(x,y)e^{-i2\pi(\frac{xu}{\lambda F}+\frac{yv}{\lambda F})}dxdy\bigg|^2 \right\rangle$$
$$(B.6)$$

If we multiply out the terms in the modulus squared bracket and use the ensemble averaged operation and the assumptions in equation (B.3) to yield:

$$PSF(u,v) \quad \frac{1}{\lambda^2 F^2}\bigg|_{aper} A(x,y)e^{i\frac{2\pi}{\lambda}W_d(x,y)} e^{-i2\pi(\frac{xu}{\lambda F}+\frac{yv}{\lambda F})}dxdy\bigg|^2 + \frac{16\pi^2}{\lambda^4 F^2}\left\langle\bigg|_{aper} A(x,y)e^{i\frac{2\pi}{\lambda}W_d(x,y)}S(x,y)e^{-i2\pi(\frac{xu}{\lambda F}+\frac{yv}{\lambda F})}dxdy\bigg|^2\right\rangle$$
$$(B.7)$$

There are a number of things to note about equation (B.7). The first is that, due to the small angle approximation, equation (B.7) no longer rigoursly conserves energy. This will shortly be corrected for via the introduction of the Strehl ratio. The first term in equation (B.7) is the classical aberrated diffraction point spread function and the second term is the first order approximation to the scattering term. Note that the scattering term scales as $\frac{1}{\lambda^4}$ and the diffraction term scales as $\frac{1}{\lambda^2}$, thus, the scattering term quickly dominates as one tends towards shorter wavelengths. It can be shown that attenuation of the first term via the Strehl ratio rigoursly conserves energy, thus equation (B.7) becomes:

$$PSF(u,v) = e^{-\frac{16\pi^2}{\lambda^2}\sigma_s^2}PSF_d(u,v) + \frac{16\pi^2}{\lambda^4 F^2}\left\langle\bigg|_{aper} A(x,y)e^{i\frac{2\pi}{\lambda}W_d(x,y)}S(x,y)e^{-i2\pi(\frac{xu}{\lambda F}+\frac{yv}{\lambda F})}dxdy\bigg|^2\right\rangle$$
$$(B.8)$$

Equation (B.8) is the sum of 2 terms: the Strehl attenuated diffraction PSF, and the linearized scattering term.

This approximation gives alot of insight into the scatter, since the scatter PSF becomes the sum of 2 terms, the first of which is the Strehl attenuated diffraction PSF and the second is proportional to the power spectral density (PSD) of the surface. If one approximates the surface PSD by a polynomial of the form:

Approach 2: Statistical Scattering Theory

 [These results will be inserted in a later document]

Approach 3: Diffraction Scattering Theory

[These results will be inserted in a later document]

Appendix C

Derivation of Extended Object Phase Diversity Algorithm

For the 2 channel phase diversity system the focal plane image in channel 1 is given by:

$$d_1(x,y) = P_1(x,y) ** O(x,y) + \eta_1(x,y) \qquad\qquad (C.1)$$

and for channel 2 by:

$$d_2(x,y) = P_2(x,y) ** O(x,y) + \eta_2(x,y) \qquad\qquad (C.2)$$

where the object **O** is the same for both channels and the point spread functions, $\mathbf{P}_1$ and $\mathbf{P}_2$ differ for each of the bands.   Note that $**$ denotes convolution. An additive noise model is assumed in this case where η_1 and η_1 are the channel 1 and 2 noise vectors respectively. Both channels ìseeî the same wavefront due to the optical systems errors, however a wavefront diversity, δW, is introduced into the system via deliberately looking at the images at different foci, and/or known differences in aberrations.   Focus error is only one form of phase diversity.   The point spread functions (PSF), for a single plane diffraction model, are given by:

$$P_1(x,y;W) = \frac{1}{\lambda^2 F^2}\left| \quad A(u,v)e^{ikW(u,v)}e^{-i\frac{2\pi}{\lambda F}(xu+yv)}\right|^2 \qquad\qquad (C.3)$$

$$P_2(x,y;W+\delta W) = \frac{1}{\lambda^2 F^2}\left| \quad A(u,v)e^{ik(W(u,v)+\delta W(u,v))}e^{-i\frac{2\pi}{\lambda F}(xu+yv)}\right|^2 \qquad\qquad (C.4)$$

Where λ is the wavelength and F the system focal length, A(u,v) is the aperture function, W(u,v) the combined atmosphere and telescope wavefront , $\delta W(u,v)$ is the diversity wavefront, u,v are the exit pupil coordinates and x,y are the respective focal plane coordinates. The goal in phase diversity is to simultaneously estimate both the wavefront W(u,v) and the object O(x,y) from a single dataset containing multiple images each with a different diversity. Figure ??? shows the phase diversity method used here in a graphical flowchart form to be discussed in more detail here.

One starts by assuming an initial guess for the wavefront W(u,v) and evaluating equations (C.3) and (C.4) via fast Fourier transform (FFT) techniques. An initial object estimate is then made, in the Fourier domain, via a Wiener filter of the form:

$$\tilde{O}(f_x, f_y) = \frac{\tilde{P}_1^*(f_x, f_y) d_1(f_x, f_y) + \tilde{P}_2^*(f_x, f_y) d_2(f_x, f_y)}{\left|\tilde{P}_1^*(f_x, f_y)\right|^2 + \left|\tilde{P}_2^*(f_x, f_y)\right|^2 + \beta}$$ (C.5)

The object spectrum is inverse FFTíd and the positivity constraint is enforced, i.e. any points which are less than zero in the object are set to zero. The PSFs are then re-estimated from the positively constrained object via Wiener filtering:

$$\tilde{P}_1(f_x, f_y) = \frac{\tilde{O}^*(f_x, f_y) d_1(f_x, f_y)}{\left|\tilde{O}^*(f_x, f_y)\right|^2 + \beta} \quad \text{and} \quad \tilde{P}_2(f_x, f_y) = \frac{\tilde{O}^*(f_x, f_y) d_2(f_x, f_y)}{\left|\tilde{O}^*(f_x, f_y)\right|^2 + \beta}$$ (C.6)

The PSFs are inverse FFTíd and again the positivity constraint applied. An iterative transform algorithm (ITA) is then used to back propagate from the respective focal planes to the pupil plane; the focal plane phase used for the ITA is from the previous iteration. A value of zero is used as a starting point. In the pupil plane the phase diversity is removed and the wavefronts averaged to yield an updated value for W(u,v) and then the entire process repeats. Typically it takes ~100 iterations before a solution is reached.

Algorithm

Define a log-likelihood function of the form (assumes Gaussian statistics):

$$E^{(j)} = \left[\mathbf{d}^{(j)} - \mathbf{P}^{(j)}\mathbf{O}\right]^T \mathbf{C}_v^{-1} \left[\mathbf{d}^{(j)} - \mathbf{P}^{(j)}\mathbf{O}\right]$$ (C.7)

where the superscripted j refers to the j-th data set and PSF realization. Note that ìjî refers to the time index. The total error metric is thus given by:

$$E = \sum_j E^{(j)} = \sum_j \left[\mathbf{d}^{(j)} - \mathbf{P}^{(j)}\mathbf{O}\right]^T \mathbf{C}_v^{-1}\left[\mathbf{d}^{(j)} - \mathbf{P}^{(j)}\mathbf{O}\right] \qquad (C.8)$$

where the noise covariance matrix is given by $\mathbf{C}_v = <\mathbf{n}\mathbf{n}^T>$, temporarily assumed stationary and uncorrelated (simplifies notation). Solving equation (C.8) for it's minimum, in a least squares sense, yields an estimator for the object given by:

$$\hat{\mathbf{O}} = \left[\sum_j \mathbf{P}^{T^{(j)}}\mathbf{C}_v^{-1}\mathbf{P}^{(j)}\right]^{-1}\sum_j \mathbf{P}^{T^{(j)}}\mathbf{C}_v^{-1}\mathbf{d}^{(j)} \qquad (C.9)$$

or if we operate only within the ìisoplanatic patchî (i.e. spatially stationary PSF) we can work in the Fourier domain and equation (C.9) becomes:

$$\tilde{O}(f) = \frac{\sum_j \tilde{P}^*_{A,j}(f)\tilde{d}_{A,k}(f) + \sum_j \tilde{P}^*_{B,j}(f)\tilde{d}_{B,j}(f)}{\sum_j \left|\tilde{P}_{A,j}(f)\right|^2 + \sum_j \left|\tilde{P}_{B,j}(f)\right|^2 + \alpha^2} \qquad (C.10)$$

This takes the form of a multi-channel Wiener filter, where is a classical Wiener filter type regularization term related to the noise spectrum. Note, a number of other regularization schemes are available. A Pseudo-Code for this algorithm is shown below for completeness.

Pseudo-Code

(1) Fourier transform the data set: $\left\{\tilde{d}_A(f), \tilde{d}_B(f)\right\}_{j=1,...,N}$

(2) Assume start set of N starting wavefronts: $\left\{w^{(0)}(u,v;t)\right\}_{j=1,...,N}$

(3) Generate set of PSFs from eqns (7) and (8): $\left\{P_A^{(p)}(x,y), P_B^{(p)}(x,y)\right\}_{j=1,...,N}$

(4) Fourier transform set of PSFs to OTFs: $\left\{\tilde{P}_A^{(p)}(f), \tilde{P}_B^{(p)}(f)\right\}_{j=1,...,N}$

(5) Evaluate object spectrum via equation (12): $\tilde{O}^{(p)}(f)$

(6) Inverse FT object spectrum: $O^{(p)}(x,y)$

(7) Object projection: $if \quad O^{(p)}(x,y) < 0.0$
$then \quad O^{(p)}(x,y) = 0.0$

(8) Estimate set of OTFs:

$$\tilde{P}_{A,B}^{\ (p)}(f) = \frac{\tilde{O}^{*(p)}\tilde{d}_{A,B}(f)}{\left|\tilde{O}^{(p)}(f)\right|^2 + \beta^2}_{\ j=1,\dots,N}$$

(9) Estimate PSFs as inverse FT of OTFs:

$$\left\{P_A^{(p)}(x,y), P_B^{(p)}(x,y)\right\}_{j=1,\dots,N}$$

(10) PSF projection:

$$if \quad P_{A,B}^{\ (p)}(x,y) < 0.0$$
$$then \quad P_{A,B}^{\ (p)}(x,y) = 0.0$$
$$for\ any\ j = 1,\dots,N$$

(11) Estimate pupil wavefronts:

$$\left\{w^{(p)}(u,v)\right\}_{j=1,\dots,N} \% \, 2\pi$$

use ITA from focal plane $\{\phi^{(p-1)}(u,v)\}_{j=1,\dots,N}$
to pupil with focal plane phase from last iteration.

(12) Apply pupil plane support constraint: A(u,v)

(13) Unwrap pupil plane phases:

$$\left\{w^{(p)}(u,v)\right\}_{j=1,\dots,N}$$

(14) Remove wavefront diversities and average residuals

(15) add diversities to residuals to get p+1 set of wavefronts

(16) Goto step (3) and repeat cycle